(Founded in 1935 by Carl Murchison)

The Journal of PSYCHOLOGY

The General Field of Psychology

EDITED BY
CARL MURCHISON

GORDON ALLPORT
Harvard University

ARTHUR G. BILLS
University of Cincinnati

E. A. BOTT
University of Toronto

LEONARD CARMICHAEL
Smithsonian Institution

HERBERT S. CONRAD
U. S. Office of Education

ELMER K. CULLER
University of Rochester

JOHN F. DASHIELL
University of North Carolina

ROY M. DORCUS
University of California at Los Angeles

HORACE B. ENGLISH
Ohio State University

FRANKLIN FEARING
University of California at Los Angeles

FRANK A. GELDARD
University of Virginia

CLARENCE H. GRAHAM
Columbia University

ERNEST R. HILGARD
Stanford University

HEINRICH KLÜVER
University of Chicago

CARNEY LANDIS
Columbia University

NORMAN L. MUNN
Bowdoin University

JOHN PAUL NAFE
Washington University

DONALD G. PATERSON
University of Minnesota

LEE EDWARD TRAVIS
University of Southern California

ROBERT C. TRYON
University of California

MICHAEL J. ZIGLER
Wellesley College

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold under the restrictions of a bilateral contract, and may not be resold for a period of five years from date of publication.

VOLUME 47

1959

Copyright, 1959, by The Journal Press
Published Quarterly by The Journal Press
Provincetown, Massachusetts, U. S. A.

AUTHOR INDEX

Adair, Harley J. 289
Altus, William D. 37
Bartley, S. Howard 289
Beatty, Frank S. 267
Beer, Michael 49
Beigel, Hugo G. 171
Buckhout, Robert 49
Burnstein, Eugene 81
Clausen, Johs. 189
Comalli, Peter E., Jr. 259
Crawford, John L. 87
Dameron, Lawrence E. 267
Deignan, Frank J. 223
Donaldson, J. 163
Dorfman, Donald 81
Ekman, Gösta E. 343
Freeman, G. L. 137
Gough, Philip B. 67
Gray, Florence E. 247
Greene, Joel E. 267
Gustav, Alice 99
Hanson, John A. 247
Horowitz, Milton W. 49
Hull, Chester D. 219
Jones, Frank Pierce 247
Joshi, Bhuwan Lal 3
Karon, Edward S. 231
Karon, Bertram P. 231
Karrer, Rathe 189
Knehr, Charles A. 297
Kohl, Richard N. 297
Kunnapas, Theodor M. 41, 305
Landes, Judah 67
Levy, Seymour 49
Magnuson, K. 163
Mayzner, M. S. 107, 117, 337
McHugh, L. 163
Morant, Ricardo B. 25
Mount, George E. 219
Nash, Harvey 277
Niner, R. 163
Noble, Clyde E. 331
O'Connell, D. N. 247
Perry, Daniel J. 219
Polin, A. Terrence 235
Prelinger, Ernst 13
Roehrig, William C. 317
Shapiro, David 127
Tagiuri, Renato 127
Tresselt, M. E. 107, 117, 337
Vanderplas, James M. 87
Walters, Richard H. 199, 209
Wapner, Seymour 259
Warren, J. M. 3, 9
Watt, F. 163
Weitzenhoffer, Andre M. 67
Werner, Heinz 259
Williams, B. 163
Zaks, Misha S. 199, 209
Ziller, Robert C. 57
Zink, A. 163

Bur[illegible] Edu[illegible] Research
Dated 3.9.70
Acc. No. J-327

TABLE OF CONTENTS

Discrimination of ambivalent cue stimuli by cats 3
By Bhuwan Lal Joshi and J. M. Warren

Perseverative reactions in chicks and kittens 9
By J. M. Warren

Extension and structure of the self 13
By Ernst Prelinger

The visual perception of the median plane as influenced by labyrinthian stimulation 25
By Ricardo B. Morant

Sexual rôle, the short story, and the writer 37
By William D. Altus

The vertical-horizontal illusion in artificial visual fields 41
By Theodor M. Kunnapas

Some perceived properties of the difference between leaders and non-leaders . 49
By Michael Beer, Robert Buckhout, Milton W. Horowitz, and Seymour Levy

Leader acceptance of responsibility for group action under conditions of uncertainty and risk 57
By Robert C. Ziller

A study of the Braid effect: (Hypnosis by visual fixation) 67
By Andre M. Weitzenhoffer, Philip B. Gough, and Judah Landes

Some effects of meprobamate on human learning 81
By Eugene Burnstein and Donald Dorfman

An experiment on the mediation of transfer in paired-associate learning . . 87
By John L. Crawford and James M. Vanderplas

Field work in the undergraduate psychology curriculum 99
By Alice Gustav

Shifts in connotative meaning of words as a function of varying amounts of previous restrictive experience 107
By M. S. Mayzner and M. E. Tresselt

Anagram solution times: A function of transition probabilities 117
By M. S. Mayzner and M. E. Tresselt

Sex differences in inferring personality traits 127
By David Shapiro and Renato Tagiuri

Pre-retirant test scores in various hobby groups 137
By G. L. Freeman

Psychological aspects of confinement in fallout shelters 163
By J. Donaldson, K. Magnuson, L. McHugh, R. Niner, F. Watt, B. Williams, and A. Zink

Mental processes during the production of dreams 171
By Hugo G. Beigel

Phosphene threshold as related to age and sex 189
By Johs. Clausen and Rathe Karrer

First steps in the construction of a scale for the measurement of aggression 199
By Misha S. Zaks and Richard H. Walters

Validation studies of an aggression scale 209
By Richard H. Walters and Misha S. Zaks

The effect of varying intramuscular dosages of atropine and banthine on the galvanic skin response 219
By Daniel J. Perry, George E. Mount, and Chester D. Hull

Two-year changes on the *ACE* by students in a college of art and architecture 223
By Frank J. Deignan

Differentiation and differential counsel of some neurotic personalities . . 231
By Edward S. Karon and Bertram P. Karon

The effects of flooding and physical suppression as extinction techniques on an anxiety motivated avoidance locomotor response 235
By A. Terrence Polin

An experimental study of the effect of head balance on patterns of posture and movement in man 247
By Frank Pierce Jones, Florence E. Gray, John A. Hanson, and D. N. O'Connell

Perception of verticality in middle and old age 259
By Peter E. Comalli, Jr., Seymour Wapner, and Heinz Werner

An investigation of the effects of reward and punishment on visual perception 267
By Frank S. Beatty, Lawrence E. Dameron, and Joel E. Greene

The behavioral world 277
By Harvey Nash

Comparisons of phenomenal distance in photographs of various sizes . . 289
By S. Howard Bartley and Harley J. Adair

MMPI screening of entering medical students 297
By Charles A. Knehr and Richard N. Kohl

Visual field and subjective center of a diamond 305
By Theodor M. Kunnapas

The influence of area on the critical flicker-fusion threshold 317
By William C. Roehrig

Ability vs. practice in paired-associate learning 331
By Clyde E. Noble

A further study in the consistency of judgments in categorizing verbal material 337
By M. E. Tresselt and M. S. Mayzner

Weber's law and related functions 343
By Gösta E. Ekman

Published as a separate and in *The Journal of Psychology*, 1959, **47**, 3-7.

DISCRIMINATION OF AMBIVALENT CUE STIMULI BY CATS*[1]

Departments of Psychology, University of California and Stanford University

BHUWAN LAL JOSHI AND J. M. WARREN

A. PURPOSE

The purposes of this experiment were to determine whether cats can discriminate stimuli presenting ambivalent cues, and to compare the performance of cats with that of monkeys tested on the same problem under similar conditions (4).

B. METHOD

1. *Subjects*

Of the six cats studied, two (Nos. 1 and 4) had approximately a year's experience on learning set (6), and oddity (1) problems. The third *S*, No. 27, had been tested for about nine months on a variety of problems, including discrimination (5) and discrimination reversal tasks. Cat No. 38 had previously been tested on discrimination and discrimination reversal for two months. The remaining two *S*s, Nos. 51 and 52, had not been trained in any learning experiments. Cats with widely different testing histories were selected for this experiment in order to ascertain whether extensive training on other problems would facilitate solution of the ambivalent cue problem.

2. *Apparatus*

Training was conducted in the Wisconsin General Test Apparatus (2, Figure 1). A grey form board containing food wells 12″ apart was used for all tests.

3. *Stimuli*

The four test objects were: large white (W) large black (B), small white (w), and small black (b); these were blocks cut from ¾″ lumber. The small objects were 4 by 4 cm. squares, and the large objects were 4 by 20 cm. rectangles. The four stimuli were paired in the following con-

---

*Received in the Editorial Office on September 5, 1958, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This research was supported by Grant M-835 from the National Institute of Mental Health, U. S. Public Health Service.

Bureau Ednl. Psy. Research
DAVID HARE TRAINING COLLEGE
Dated........................
Accs. No......................

figurations (+ and — designate the rewarded and unrewarded objects, respectively):

1. b+ w—
2. W+ B—
3. w+ W—
4. B+ b—

4. *Procedure*

In general, the procedures followed in this experiment were identical with those of Noer and Harlow (4): the noncorrection method was used, 25 trials were presented each day, and the criterion of solution was 45 correct responses in 50 successive trials (two days' testing).

None of the cats became emotionally disturbed during the experiment, so that intertrial feeding was not necessary, as it was with monkeys (4). Pork kidney was used as the incentive.

TABLE 1
Number of Errors Made by Each Cat on Individual Tests

Test	Configurations	Cat number 1	4	27	38	51	52
I	1	73	36	35	143	144	10
II	2	57	162	48	132	194	91
IIIa	1, 2	29	64	9	81	57	15
IIIb	1, 2	8	11	4	5	4	13
IV	3	45	36	25	34	81	28
V	4	54	60	56	57	158	35
VIa	3, 4	181	256	208	253	214	7
VIb	3, 4	7	13	23	16	20	49
VII	2, 3, 4	16	98	32	63	28	7
VIII	1, 2, 3, 4	72	36	134	16	43	18
Total		542	772	574	800	943	273

a = trials presented in alternate groups of five.
b = trials presented in predetermined irregular sequence.

Since the cats were trained according to the same stepwise procedure which was fully described by Noer and Harlow (4), it is unnecessary to reproduce their detailed treatment here. The essential steps in training are given in Table 1. It should be noted that some of Noer and Harlow's monkeys were trained on four additional configurations.

C. Results

The performance of the individual cats is summarized in Table 1, which shows the number of errors made on the various tests and over the entire series of tests. These data indicate that all of the cats solved every one of the tests and, further, that four cats learned Test VIII, which required simultaneous mastery of four antagonistic habits (b+ w—, B— W+, w+ W— and B+ b—) more quickly than the uncomplicated Test I which required only a simple brightness discrimination (b+ w—).

Total errors over the series of tests is a serviceable index of overall performance. This measure does not suggest that previous testing experience contributes materially to efficient solution of the ambivalent cue problem, since cat No. 52, an experimentally naive *S*, learned far more rapidly than did the three cats which had a year's experience with other learning tasks.

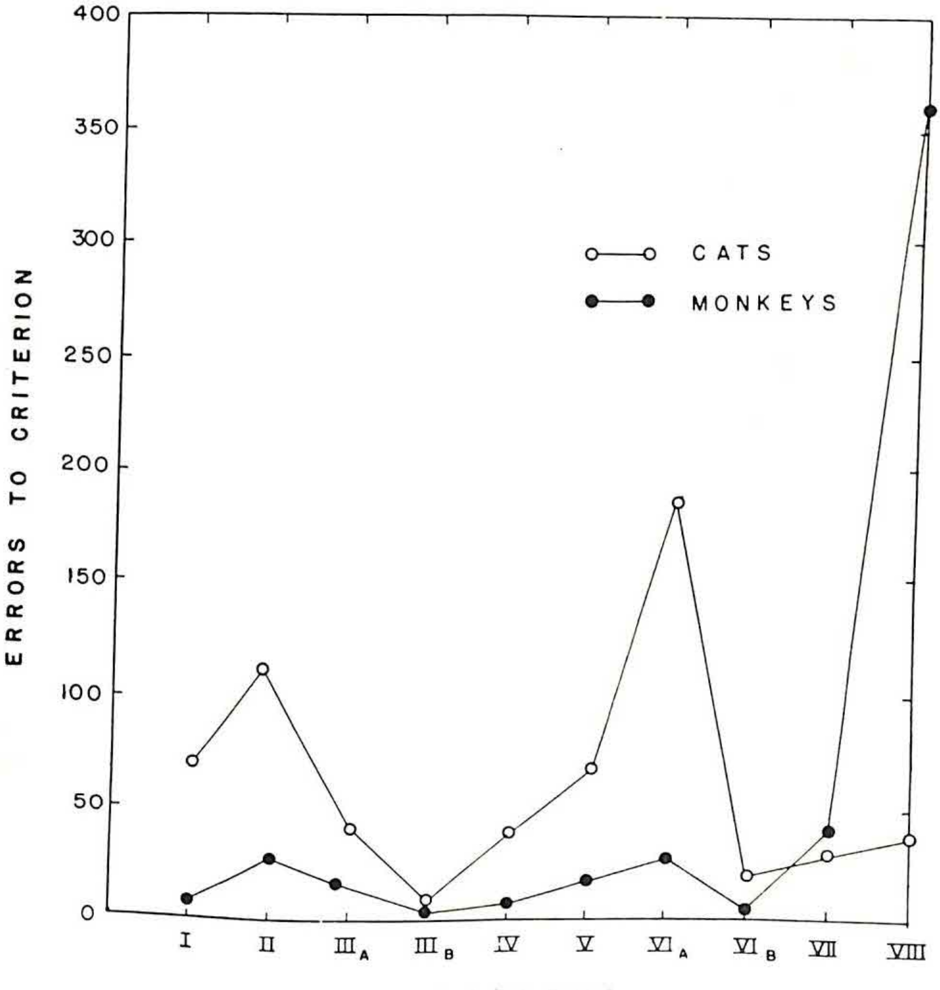


FIGURE 1

Solution of the ambivalent cue problem by cats and monkeys is compared in Figure 1, which shows mean errors to criterion by the two groups on Tests I through VI_B. The points for Tests VII and VIII are the medians for the cats and monkeys; medians were used because some of the monkeys required remedial training before passing Test VII, and a number of monkeys failed Test VIII, apparently because of emotional disruption. This figure indicates that, although the cats were generally inferior to the monkeys on Tests I through VI_B, the monkeys made almost 10 times as many errors in mastering Test VIII which required simultaneous integration of four conflicting habits. A *t*-test of the significance of the differences indicates that the performance of the monkeys was significantly inferior to that of the cats on Test VIII ($t = 3.81$, $P = .05$, with 11 *df*).

D. Discussion

The results of the present experiment are difficult to interpret. Previous research has indicated that cats are inferior to primates in the formation of learning sets (6) and in solving the oddity problem (1), yet the cats studied in this investigation were markedly superior to Noer and Harlow's monkeys on the test which required concurrent mastery of four antagonistic habits.

Previous investigators (3, 4) have recognized that the ambivalent cues problem, typified by Test VIII, is susceptible to equivocal interpretations: the *S* may learn to respond to peculiar relational properties presented by each pair of stimuli, or he may respond conditionally to the absolute properties of each stimulus in terms of multiple signs. The experimental design does not permit any unambiguous test of these hypotheses. Lacking such a definitive test of the alternative theories, it would be misleading to maintain that cats are more competent at solving complex discriminations than monkeys, since the possibility that the two species learn such tasks in different ways is not excluded. Indeed, the possibility that cats and primates responded differently to the problem is suggested by the observations that the cats remained calm, coöperative, and well-motivated throughout the experiment, while both monkeys (4) and chimpanzees (3) tested on this problem were subject to emotional and motivational disturbances.

The difficulties encountered in interpreting this experiment illustrate the risks involved in attempting to compare species' performance on a single learning task. All is well if the learning scores reflect variations in a single homologous function, but the technique becomes suspect if there are inversions in the relation between phyletic status and apparent ability to adjust to the situation, and if there is no method for precise identification of differential modes of adapting to the test.

REFERENCES

1. BOYD, B. O., & WARREN, J. M. Solution of oddity problems by cats. *J. Comp. & Physiol. Psychol.*, 1957, **50**, 258-260.
2. HARLOW, H. F. The formation of learning sets. *Psychol. Rev.*, 1949, **56**, 51-65.
3. NISSEN, H. W. Ambivalent cues in discriminative behavior of chimpanzees. *J. of Psychol.*, 1942, **14**, 3-33.
4. NOER, M. C., & HARLOW, H. F. Discrimination of ambivalent cue stimuli by macaque monkeys. *J. Gen. Psychol.*, 1946, **34**, 165-177.
5. WARREN, J. M. Solution of object and positional discriminations by cats. *J. Comp. & Physiol. Psychol.* (in press)
6. WARREN, J. M., & BARON, A. The formation of learning sets by cats. *J. Comp. & Physiol. Psychol.*, 1956, **49**, 227-231.

Department of Psychology
Stanford University
Stanford, California

Published as a separate and in *The Journal of Psychology*, 1959, **47**, 9-12.

PERSEVERATIVE REACTIONS IN CHICKS AND KITTENS*

Department of Psychology, Stanford University

J. M. Warren[1]

A. Problem

Hamilton (1, 2) found a consistent, inverse relationship between behavioral stereotypy, on an insoluble multiple choice task, and the phyletic status of several species of mammals. The performance of infra-mammalian species on this task has not been described.

In the present experiment, the performance of chicks and kittens on Hamilton's multiple choice problem was investigated. Both species were studied in a modification of Hamilton's apparatus which required only simple locomotor responses; the chicks were selected to determine the characteristics of a sample of submammalian adjustment to the problem, and the kittens to provide an interspecies comparison free from confounding of species and apparatus variables.

B. Method

1. *Subjects*

The seven White Leghorn chicks were 24 to 32 days old at the beginning of the experiment. They had previously been trained on a discrimination reversal problem, and so had become accustomed to frequent handling and to the deprivation schedule.

Four of the five five kittens studied had been tested on discrimination problems for about a month; the fifth kitten had not been trained on any prior learning task. The kittens were approximately 5 months of age.

2. *Apparatus*

The floor plan of the apparatus used in the present experiment is presented in Figure 1. The partitioning of the goal boxes prevents *S* from seeing whether a given box contains food until he passes through the second opening. This modification, the only important deviation from Hamilton's plans, eliminates the necessity for teaching the animal to push open a door or climb over a barrier as Hamilton did.

---

*Received in the Editorial Office on September 5, 1958, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This research was supported by Research Grant NSF-G3278, from the National Science Foundation.

3. *Procedure*

On any trial, only one of the four goal boxes, which are equidistant from the starting point and identical in appearance, contains food. The designation of the rewarded alternative is varied randomly from trial to trial, subject only to the constraint that the same choice is never correct on two consecutive trials. The only thing that *S* can learn, then, is to avoid, on any particular trial, the door which was correct on the immediately preceding trial.

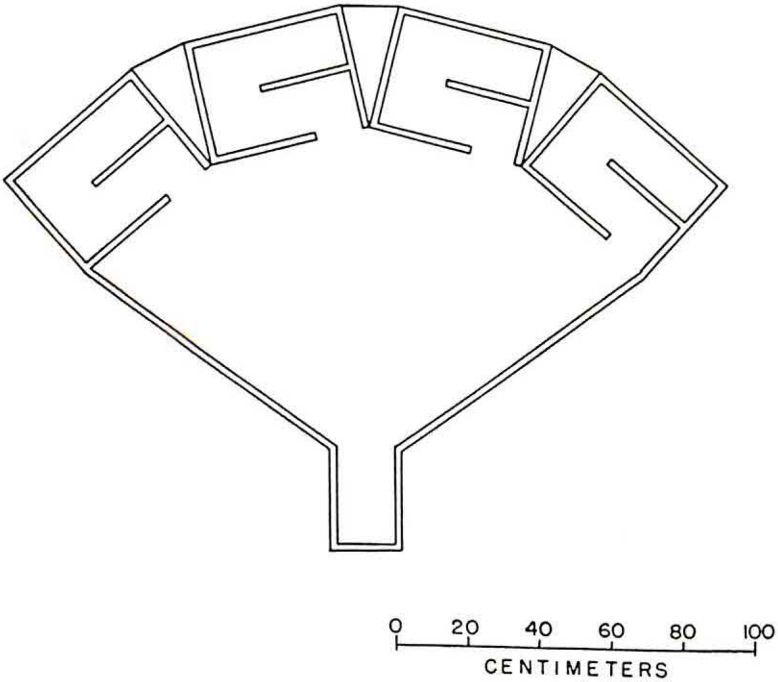


FIGURE 1
FLOOR PLAN OF THE MULTIPLE CHOICE APPARATUS

Every effort was made to duplicate Hamilton's adaptation and testing procedures and his method of analyzing the results. The essentials of his procedures are briefly: A "trial" is a sequence of responses terminated when *S* finds the food reward. An animal is tested 10 trials a day for 10 days, and performance on the series of 100 trials is described in terms of (*a*) total responses (i.e., entries into specific compartments), and (*b*) types of responses made in searching for the correct door. Trials are initially dichotomized into two types, unclassified trials being those on which fewer than three different responses are made. Those trials on which three or more responses are made are assigned to five classes representing, according to Hamilton, progressively less adequate modes of adjustment:

Type *A*: All three possible doors are tried, once each; no response is made to the impossible (just previously correct) door.

Type *B*: All four doors are tried, once each and in an irregular order.

Type *C*: All four doors are tried in a regular sequence, from one side to the other, but without any repeated responses.

Type *D*: The same door is tried more than once but the two attempts are not successive.

Type *E*: "This type includes various highly inappropriate modes" of response, those trials on which *S*: (*a*) makes two consecutive responses to the same door; (*b*) responds to two or more incorrect alternatives repetitively and in a regular order; (*c*) makes seven or more errors.

The reward for the chicks was finely cracked corn, a food preferred to the maintenance diet of starter mash. The chicks were permitted to feed *ad lib* for three hours each day, immediately after testing, and were thus tested under 21 hours food deprivation. Well nourished cats can survive for weeks without food, so no attempt was made to establish a deprivation schedule for the kittens. They had continuous access to dry dog food in their home cages, and were rewarded with kidney, a highly preferred food, during the test sessions.

Neither the chicks nor the cats were shocked in the apparatus at any time.

C. RESULTS

The results are presented in Table 1, which shows mean total errors, mean classified trials, and the distribution of reaction types for the chicks and kittens. The scores of the two groups are almost identical on every measure of performance. The kittens made 3 per cent fewer errors than the chicks; the largest discrepancy in frequency of a particular response type was 6 per cent.

TABLE 1
RESPONSES OF CHICKS AND KITTENS ON HAMILTON'S MULTIPLE CHOICE PROBLEM

	N	Mean total responses	Mean classified responses	Percentage of classified responses				
				A	B	C	D	E
Chicks	7	349	35	7	10	15	12	56
Kittens	5	340	36	4	16	11	15	54
Hamilton's kittens	2	387	52	7	11	0	14	68

The scores made by Hamilton's kittens, which are also given in Table 1, are somewhat higher than those made by the kittens studied in this experiment. These scores are included to show that the modified apparatus did not produce inferior performance by the kittens, and that the variation between small samples of kittens is greater than the difference obtained between kittens and chicks.

D. DISCUSSION

The results of the present experiment indicate that the performance of chicks and kittens on the Hamilton multiple choice problem does not differ significantly. Assuming the validity of Hamilton's *a priori* analysis of the problem, these findings would have to be interpreted as evidence of a high level of adaptive intelligence in chicks, since they "abstract," i.e., avoid making stereotyped or perseverative errors, as successfully as an intermediate mammalian species.

It is difficult to imagine a conclusion in more violent conflict with common sense. A zoölogist might convince a sceptical farmer that there is such an animal as a giraffe; a psychologist could never make him believe that a chicken is as "intelligent" as a cat.

Rather than accept as meaningful data which point to an untenable conclusion, one must question the validity of the method which yields such data. Probably the chief deficiency of the multiple choice method is the very characteristic which gives it face validity, its complexity. Error scores reflect the influence of many variables which cannot be evaluated separately with the precision attainable in simpler learning and transfer situations.

E. SUMMARY

Seven White Leghorn chicks and five kittens were tested on Hamilton's insoluble multiple choice problem. The performance of the two species did not differ significantly. The validity of the task for inter-species comparisons was questioned.

REFERENCES

1. HAMILTON, G. V. A study of trial and error reactions in mammals. *J. Anim. Behav.*, 1911, 33-66.
2. ———. A study of perseverance reactions in primates and rodents. *Behav. Monog.*, 1916, **3**, No. 13, 1-65.

Department of Psychology
Stanford University
Stanford, California

Published as a separate and in *The Journal of Psychology*, 1959, **47**, 13-23.

EXTENSION AND STRUCTURE OF THE SELF*

Department of University Health, Yale University

ERNST PRELINGER[1]

A. INTRODUCTION

Several of the current theories of personality represent an individual's psychological reality by means of a conceptual space (6, 7, 9). This is especially true for the "field" and phenomenological theories, but the "object world" of psychoanalytic theory implies a similar spatial conceptualization (4). These theories postulate within their conceptual space a relatively segregated region which corresponds to that area of experience which an individual perceives to be his own "self." The "self" has been the object of a considerable number of recent experimental studies. They have emphasized topics such as attitudes towards the self as related to adjustment, changes in self-perception as a result of therapy, or the relationships between persons' self-evaluations and the perceived evaluations of themselves by others. Little attention in any form other than introspective description has, on the other hand, been paid to the internal structure or the contents of this hypothetical self-region. The present study attempts such an investigation.

A theoretical construct, such as that of the self-region, gains in usefulness the more clearly it can be related to observed events. We first asked, therefore, if people can actually observe something in their experience which could be said to correspond to this hypothetical self-region. If they can, the internal structure of this self-region could next be investigated. That introspective observations of a "self"-region, in contrast to a "not-self" region can be made has been suggested by a number of writers, especially by William James (5). He emphasized that we tend to perceive certain objects to be parts of our selves while other objects are not. Inclusion in the self gives objects a certain quality which other, not-included objects lack.

In the present study the following questions were investigated: (*a*) Do *S*s perceive something corresponding to a self-region, and if they do, can their perceptions be made accessible to objective analysis? (*b*) If there is a perceived self-region, what can be determined concerning its internal struc-

---

*Received in the Editorial Office on September 10, 1958, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The data for this study were collected when the author was a clinical psychologist at Walter Reed Army Hospital, Washington, D. C.

ture? (*c*) What variables can be found to be related to the division of an individual's total life space into a self and a not-self area?

B. METHOD

1. *Procedure*

A set of 160 items, 20 in each of eight groups, was developed. They were originally selected from lists prepared by over 100 students at Brooklyn College who enumerated things which they believed to be parts of their own selves. Some items, especially body parts, recurred frequently in their lists, while others were mentioned only rarely. The choice of item categories to be used in the present study was influenced by James' inventory of the constituents of the "material self" and the "social self" (5, p. 292) as well as by Allport's descriptions (1, p. 217). Only items sufficiently devoid of very personal connotations were selected so that the total list could be applicable to practically anyone. All items, furthermore, had to belong clearly in one of the following eight content categories (a few randomly selected items of each group are given as examples): (*a*) Psychological or intraorganismic processes (The conscience, An itching on the sole of the foot, A feeling of being aroused sexually, The sense of vision). (*b*) Body parts (Only commonly known ones were chosen such as The skin, The genital organs, The throat, The fingers). (*c*) Objects within the close physical environment (defined as between body periphery and the walls of the rooms in which *S*s were tested; Dirt on the hands, The furniture in this room, The water touching the skin when you are bathing, The wall behind you). (*d*) Distant physical environment (extending from just outside "this" room; The adjoining room, The downtown area of this city, The basement of this building, The moon). (*e*) Personal identifying characteristics and attributes (Age, Date of birth, Occupation, The education you have had). (*f*) Possessions and productions (Watch, Perspiration on the forehead, Tears coming out of the eyes, Toilet articles). (*g*) Other people (The people in your hometown, The person in the earliest childhood memory, People who work with you, Father). (*h*) Abstract ideas (The results of science, The morals of our society, The law, The system of democracy).

The items, typed on separate cards and shuffled, were administered with approximately the following oral instructions: "If you think of all the things in the world, you may notice that they can be divided into two groups: those which are a part of your own self and those which are not. Into which group a thing falls is different for different people, but I wonder if you could give me an example of something that is definitely a part of *your* own self."

Anything named by *S* was accepted by *E*, who then continued: "And now give me an example of something which is definitely *not* a part of your own self." When this second example was obtained *E* said: "Now here I have a set of items and I want you to decide for every one of them if it is a part of your own self or if it is not. Sort the items into two piles. Do it quite quickly and on the basis of your very first feeling about each item." After the first sorting into two piles ("Part of my Self" and "Not part of my Self") *S*s were asked to sort each one of the two obtained piles into two more according to whether they felt sure the item was (was not) part of their own self or whether they felt not so sure. Therefore every *S* sorted the items into four categories which were assumed to fall onto a continuum. Scores of 0, 1, 2, or 3 were assigned to each item according to the degree to which *S* included the item into his self, the highest score representing definite inclusion. Items scored 0 and 1 were felt not to be part of his self by *S*, those scored 2 and 3 on the other hand were parts of *S*'s self. The score given to an item by *S* is designated by v, and the sum of all the v's for a *S* by *C*.

2. *Subjects*

The results of the present study are based on the responses of 60 normal enlisted soldiers, members of Troop Command at Walter Reed Army Hospital. None of them was a patient of any kind. They ranged in age from 20 to 56 years, in education from 5 to 19 years, and their intelligence scores (Army Aptitude Area I tests, standardized with a mean of 100 and an *SD* of 20) ranged from 72 to 153. Their means on these three variables were 27.7, 12.5, and 111 respectively. It should be noted that the above described instructions presented no difficulties for any, not even the less intelligent *S*s.

C. RESULTS

1. *Perception of a Self-Region and Measurement of its Extension*

The task of sorting the items was found to be quite easy by all *S*s. Although they often had difficulties in deciding if a particular item was part of their self or not, the task on the whole seemed obvious to them. That some objects are a part of one's self while others are not appeared to be a commonsense reality to all *S*s. The perception of a segregated self-region, implying the perception of a boundary between self and not-self, was therefore clearly present in the *S*s.

Wide individual differences were found among *S*s concerning the number of items felt to be included in their selves. This number ranged from 25 to 129 with a median of 83. The variability in *C*-scores similarly was considerable, as indicated by a range from 82 to 387 with a median of 250. If the number of items included in the self is used as a measure of the extension of the self-region, we may state that *S*s differ widely in the degree of such self-extension.

Can self-extension be reliably measured? In order to answer this question the internal consistency of the variable measured by *C*-scores was investigated. The sums of all the item-scores on each one of the eight groups of items (subscales) were correlated with the total *C*-scores over the 60 *S*s. The resulting coefficients as well as the highest correlations found between each one of the eight subscales and one of the remaining seven are given in Table 1. These data indicate a high degree of internal consistency within the total set of items and the intercorrelations among subgroups show that they share a considerable amount of systematic and reliable variance.

TABLE 1
ITEM GROUP COMMUNALITIES AND CORRELATIONS WITH TOTAL *C*-SCORES*

	Item groups							
Relationships	1	2	3	4	5	6	7	8
Correlation with total *C*-score	.84	.64	.78	.78	.97	.97	.93	.96
Highest correlation with any other subscale	.79	.60	.80	.82	.95	.94	.95	.90

*All coefficients significantly different from zero ($p < .01$).

The high intercorrelations among the eight subscales might suggest that the variable measured by *C*-scores approaches unidimensionality. If it does, the items should be scalable in terms of Guttmann's criteria (3). Accordingly, two samples of eight items each (*A* and *B*), gradated in about equal steps from 90 per cent to 10 per cent in terms of the frequency with which they were included into *S*s selves but otherwise picked at random, were selected. A third sample (*C*) of 10 items gradated from 80 per cent to 20 per cent was also chosen. The 60 *S*s were divided into three random samples of 20 each for purposes of replication and reproducibility coefficients were computed (Table 2). The coefficients for item sets *A* and *B* tend to

TABLE 2
REPRODUCIBILITY COEFFICIENTS OF THREE SETS OF ITEMS FOR THREE RANDOM SAMPLES OF *S*S

	Sets of items		
Subject sample	A	B	C
I	.92	.93	.83
II	.92	.90	.80
III	.84	.88	.81

be spuriously high because of the wide range of inclusion percentages represented in them. Set *C*, with inclusion percentages restricted yields more useful values in terms of recent and more rigorous criteria for scale analysis (2, p. 284) and the coefficients are, as would be expected, lower. No attempts were made to raise the reproductibility of response patterns on any of the three scales by combining response categories. Items scoring 0 or 1 on a subject were consistently considered as excluded, items scoring 2 and 3 as included into the self of an *S*.

The coefficients in the *C* column are too low to suggest acceptable unidimensionality and scalability of these items. Nevertheless the data so far point toward the existence of a variable of self-extension which can be measured with reasonable reliability at least by means of the presently used items and methodology. Question 1 as stated at the outset of this paper therefore seems to have been answered in the affirmative.

2. *The Internal Structure of the Self*

The eight groups of items, each one representing a specific content category, were selected in order to test a hypothesized internal structure of the self. William James' statements served as a basis for the hypothesis that different groups of items would be included into an "average" self with different degrees of frequency and intensity (5, p. 292). Table 3 gives the mean v-values of items within every one of the eight groups for each of three samples of 20 *S*s. The stability of the obtained values is indicated by their similarity for the three samples of *S*s. The significance of the variance between the means of item groups was tested by comparing it with the variance of scores within the item groups. Computed separately for the three samples of *S*s, *F*-ratios of 44.90, 49.41, and 60.58 were obtained, all significant well beyond the .001 level (for *df*'s of 7 and 152). A certain degree of homogeneity of the item groups is thus suggested, which seems to indicate that the a priorily established item groups had internal coherence also in the *S*s' perceptions. The results indicate further an apparently very stable and sig-

TABLE 3
MEAN v-VALUES FOR EIGHT ITEM GROUPS FOR THREE RANDOM SAMPLES OF *S*S

Subject sample	Item groups								
	1	2	3	4	5	6	7	8	Overall mean
I	2.48	2.98	.63	.19	2.16	1.51	1.10	1.33	1.55
II	2.41	2.98	.73	.15	2.20	1.69	1.10	1.28	1.57
III	2.48	2.97	.55	.24	2.29	1.52	1.11	1.47	1.58
All samples	2.46	2.98	.64	.19	2.22	1.57	1.10	1.36	1.57

nificant rankorder among item groups. Body parts are considered to be parts of the self with greatest frequency and certainty. This would seem to be consistent with the assumption that the original differentiation between self and not-self is greatly influenced by sensory feedback obtained by the small child when manipulating parts of his body but lacking when external objects are manipulated. Thus the body is frequently considered to be the "core" of the self. Then follow Psychological and intra-organismic processes, Personal identifying characteristics, and Possessions and productions. The mean v-values of these scales all lie above 1.50, which is the hypothetical borderline between self and not-self as implied by the instructions given and by the scoring system. With mean v-values falling into the "not-self" region follow in descending order: Abstract ideas, Other people, Close physical environment, and Distant physical environment. On the whole it seems that items whose v is higher than 1.50 typically are those which can be defined socially only with reference to a person (body, possessions, name), while for lower scoring items (The system of democracy, The moon) such a definition is not necessary or even possible.

The reliability of the results suggesting the existence of a rankorder among items in terms of the degree with which they are perceived to be part of the self was checked once more by comparing the rankorders of the single items in terms of inclusion into the self among the three samples of *S*s. The 20 body part items were omitted from this comparison, since they practically all tied for the first rank. For the remaining 140 items rho's were computed between the three samples. They were for samples I and II: .97, I and III: .96, II and III: .95. All these values are significant well beyond the .01 level.

3. *Variables Associated with the Inclusion of Items into the Self*

An initial attempt was made to investigate some factors which might be assumed to be related to the different degrees with which items were included into their selves by *S*s. One hypothesis was derived from Bruner's discussion of a self-potency variable and is also related to Rogers's thinking (both quoted in 8, p. 539). It was expanded by taking into account the capability of objects to affect a person, and states: Objects over which a *S* has control or which he can manipulate, as well as objects by which he can be affected in one fashion or another, are more likely to be perceived as parts of his self than items for which neither is the case. A group of five judges was asked to sort all 160 items into three groups: those which in general are predominantly under the control of people, those by which people in general are primarily affected, and those for which neither is true, which are predominantly neutral. All five judges agreed completely on the placement of 44 items, four of them agreed completely on 98 items. The mean v-values for the three groups of items were computed and are shown in Table 4. The mean v-values of "control" and "affected-by" items are not significantly different and indicate that both kinds of items tend to be included into the self (both values greater than 1.50). The "neutral" items on the other hand tend to be highly significantly less included into the self ($p <.001$).

TABLE 4

Mean v-values of "Control," "Affected-by," and "Neutral" Items

	Control	Affected-by	Neutral
44 items (Five judges agreed)	1.94	1.84	.13
98 items (Four judges agreed)	1.95	1.83	.17

The eight content categories of items show some notable differences with regard to the factors of "control," "affected-by," and "neutrality" (Table 5).

TABLE 5

Mean Judgments of Degree of "Control," "Affected-by," and "Neutrality" for Eight Item Groups

	Item groups								
Item categories	1	2	3	4	5	6	7	8	Total scale
Control	2.10	2.25	2.35	.05	1.75	2.30	.75	1.60	1.64
Affected-by	2.80	2.45	.80	.50	2.50	1.70	3.40	2.50	2.08
Neutral	.10	.30	1.85	4.45	.75	1.00	.85	.90	1.28

The three values for each item group add up to five, which is the number of judges. A comparison of the amount of "neutrality" of each item group with its mean v-value (Table 3) shows a high, inverse relationship between the two sets of values (rho $= -.88$, $p < .01$). The hypothesis that items which are under our control or which affect us are more likely to be included into the self than neutral items seems thus well supported.

Another variable perhaps related to the degree with which items are included into the self is physical distance from the body. It was hypothesized that if such a relationship exists it might be found most clearly among the items in Group 3 (objects in the close physical environment, defined as the space between the skin and the inside walls of "this" room). The 20 items in this group were rated by seven judges, different from the five just mentioned before, on a physical distance scale using the following steps: Score 5: Penetrating the body periphery, 4: Directly contacting major area of the body surface, 3: Directly contacting minor area of the body surface, 2: Up to two inches distant from body surface, 1: More than two inches distant from body surface. The seven distance ratings for each item were averaged and the 20 resulting averages were ranked. These ranks were correlated with the ranked mean v-values of the items and a rho of .79, $p < .01$ was obtained. Physical distance of an item then seems to be related to the degree with which it is included into the self at least within this group of items. It would seem likely, however, that this relationship holds only as long as other variables such as emotional values of items, etc., are held reasonably constant. The "neutrality" of the items in this group indeed was relatively high as Table 5 shows.

An attempt was further made to relate the degree of inclusion of items to those areas customarily defined as id, ego, and superego. Five psychiatrists who had not served as judges in any other part of this study were asked to judge the 20 items of the first group (Psychological processes) in terms of which of the following categories they are primarily representing: id, ego, superego, somatic processes. These judges were second and third year psychiatric residents at Walter Reed. At least three of them agreed on the placement of 17 items: *Id*: A dream, Extreme danger, A feeling of being aroused sexually, Intense hunger; *Ego*: Thoughts, Worry about a difficult decision, A feeling of pride, Judgment, Control over emotions; *Superego*: A feeling of anxiety, An unpleasant memory, The conscience; *Body processes*: An itching on the sole of the foot, A headache, The sense of hearing, The sense of touch, The sense of vision. Mean v-values were computed for each of the four categories. Analysis showed variance associated with these four categories to a degree which just misses the .05 level. However, some of the body items, especially the three sensory ones, could perhaps be considered primary autonomous ego factors in Hartmann's sense (4). So, for purposes of further investigation, they were combined with the ego items. Id and superego items also were combined into a "non-ego" category. If the mean v-values of these two groups of items were compared it was found that the "ego"-items had on the whole greater "self-quality" than the "non-ego" items (Table 6). Computed separately for the three random groups of subjects the significance levels using two-tailed tests were $<.05$, $<.01$, $<.10$, respectively. The procedure of combining categories, however, is somewhat arbitrary and the just described results are therefore nothing more than preliminary hints.

It is interesting to compare the "ego" and the "non-ego" items in terms of their standing in the "control," "affected by," and "neutral" categories (Table 6). It is important to note that the judgments involved in the two

TABLE 6
AVERAGE "CONTROL," "AFFECTED-BY," "NEUTRALITY," AND MEAN v-VALUES OF EGO AND NON-EGO ITEMS

	Control	Affected-by	Neutral	Mean v
Ego ($N=10$)	3.8	1.2	0.0	2.65
Non-ego ($N=7$)	1.3	3.5	0.2	2.23

categorization procedures were made by different judges. Ego items tend to be judged to be more under control, while non-ego items are judged more frequently to be agents by which "one" is affected in some way.

Evidence concerning the relationship of inclusion vs. exclusion of items into the self or of C-scores with other variables was inconclusive. Education and intelligence are positively related to scores on all eight item groups as well as to C, but only a few of the coefficients approach the .05 level of confidence, while none passes it.

D. DISCUSSION

Some important limitations of the findings of this study have to be mentioned. First of all, the results are specific to the group of *S*s and to the items used. The standing of the *S*s within any larger part of the population can be relatively clearly specified by their known ages, occupation, sex, educational level, and intelligence. On the other hand it is very unclear from which universe the items are drawn. The universe was obviously delimited by instructing the college students to write lists of items which were part of their selves and some of the delimiting factors can be hypothesized: They would probably include the capability of an item to be expressed in words, the possibility of perceiving some relation between the item and the "self," certain cultural definitions of the self, etc. However, at the present time these delimitations are not sufficient to define the universe.

The extension of the self-region of a given *S* in this study is defined in terms of items included or excluded. Thus, the self-quality is not studied independently, but presupposes perceptions of objects. Its extension if studied by the present method then becomes a function among other things of the objects experienced by *S*s. The finding that all *S*s included all body part items into their selves, for instance, may be partly due to the fact that only familiar organs appeared among them. It might have been different had unfamiliar items (the labyrinth, the pancreas, etc.) been included. Here again the findings would probably have been influenced by the presence or absence in *S*s of the knowledge that these items represented parts of the body.

8-327

Do such considerations lead to the conclusion that all the reported findings are relative and dependent on too many vague and ill-defined variables? The finding of a very stable and quite meaningful rankorder among items in terms of inclusion into the self and the described relationships with outside variables speak against such a conclusion. Furthermore, the rankorder of item groups as evident from Table 3 is in fair agreement with the hypothetical hierarchies formulated by William James, Allport, and Bruner. McClelland also arranges categories of objects along a theoretical "distance from the self" continuum (8, p. 539). The following categories are coordinated by him to five steps of increasing distance: (*a*) Me, Free will, (*b*) My body, My conscience, (*c*) My belongings, (*d*) My friends, (*e*) Strangers, Physical universe, etc. Mean v-values for those items from the present study which presumably correspond most closely to McClelland's categories were determined for all of his five steps except the first, for which no appropriate items were available. The values were for (*b*): 2.97, (*c*): 1.35, (*d*): .60, (*e*): .43. Thus the findings of this study seem to be reasonably well related to the still very loose body of theory concerning itself with the internal structure of the self.

One further limitation must be mentioned. While the obtained results show a certain internal coherence they are not at this time in any demonstrated relationships to behavioral data, to other established personality constructs, or to the processes of adaptation and adjustment in general. Further studies are on the way, however, to investigate some of these relationships.

E. Summary

An investigation of the phenomenon of a segregated "self-region" as contrasted with a "not-self-region" was attempted. *S*s had to sort a set of items, representing eight different content categories, into those which they felt were parts of their selves and into those which they felt were not. The following findings were obtained:

1. *S*s had no difficulty in determining whether an item was a part of their own selves or not. Thus they seemed to have rather clear perceptions of a segregated self-region.

2. *S*s differed widely in the number of items "included" into their selves, and it was shown that this variable of "self-extension" has high internal consistency.

3. Item groups representing different content were found to be relatively homogeneous and to show a very definite and meaningful rankorder among themselves in terms of the degree to which they were included into *S*s' selves.

4. Among variables related to the degree to which items are included into *Ss'* selves were found to be the degree to which items are in general capable of being controlled by people or are capable of affecting them. Physical distance was found to be another important factor. An attempt was also made to relate the degree of "self-ness" of items to their categorization in terms of id, ego, and superego.

Some relationships of the findings to theoretical statements made by other writers were discussed.

REFERENCES

1. ALLPORT, G. W. Personality. New York: Holt, 1937.
2. FORD, R. N. A rapid scoring procedure for scaling attitude questions. In: (Riley, M. W., Riley, J. W., & Toby, J., *Eds.*) *Sociological Studies in Scale Analysis.* New Brunswick: Rutgers Univ. Press, 1954.
3. GUTTMAN, L. The basis for scalogram analysis. In (Stouffer, S. S., et al.) *Measurement and Prediction, Studies in Social Psychology in World War II.* (Vol. 4.) Princeton: Princeton Univ. Press, 1950.
4. HARTMANN, H. Comments on the psychoanalytic theory of the ego. *Psychoanal. Stud. Child,* 1950, **5**, 74-96.
5. JAMES, W. The Principles of Psychology. New York: Holt, 1890.
6. KOFFKA, K. Principles of Gestalt Psychology. New York: Harcourt, Brace, 1935.
7. LEWIN, K. A dynamic theory of personality. New York: McGraw-Hill, 1935.
8. MCCLELLAND, D. Personality. New York: Sloane, 1951.
9. SNYGG, D., & COMBS, A. W. Individual Behavior. New York: Harper, 1949.

Division of Student Mental Hygiene
Department of Health
Yale University
109 College Street
New Haven, Connecticut

Published as a separate and in *The Journal of Psychology*, 1959, **47**, 25-35.

THE VISUAL PERCEPTION OF THE MEDIAN PLANE AS INFLUENCED BY LABYRINTHIAN STIMULATION*

Department of Psychology, Brandeis University

RICARDO B. MORANT[1]

A. INTRODUCTION

Following Flouren's classic work on the sectioning of the semi-circular canals, various physiologists have shown the interrelationship between labyrinthian stimulation and the tonicity of the musculature (2, 6, 7). Sherrington (17), for example, discusses the similarity between stimulation of the labyrinthian receptors and other proprioceptor organs of the body in that both are the source of tonic reflexes. Since labyrinthian stimulation affects body tonicity, we might expect from theories which include tonic factors as essential constituents in the organization of perception, such as the sensory-tonic theory (20), that it would also affect space perception. The results of a previous experiment dealing with one cardinal spatial dimension, verticality, conformed with this expectation (19). The present experiment was designed to study the influence of labyrinthian stimulation on the visual perception of another cardinal dimension of personal space, the median plane.

The median plane is objectively defined as the extension of the median sagittal plane of the body and, subjectively, as straight-ahead. Normally, the objective and subjective median planes are identical. However, under certain conditions (3, 4, 18, 22), the physical position perceived as straight-ahead does not coincide with the objective median plane. We must, consequently, distinguish between the position experienced as straight-ahead and the position which is physically straight-ahead; that is, respectively, between the apparent median plane and the objective median plane.

We must emphasize that it is the possible change in the position of the subjective median plane which we are investigating and not other visual phenomena such as illusory forms of motion, often reported to occur under conditions of labyrinthian stimulation.[2] In the present experiment, labyrinthian stimulation was induced by acceleratory and deceleratory rotation of *S* around his vertical axis. The position of the apparent median plane was inferred from the position to which a spot was adjusted by *S* when required to place it straight-ahead.

---

*Received in the Editorial Office on September 16, 1958, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] Part of the study reported below has previously been presented in a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Clark University in 1952. This research was supported by Grant M-348 from the National Institute of Mental Health, of the National Institute of Health, Public Health Service. The author wishes to acknowledge his debt to the directors of the grant, Drs. H. Werner and S. Wapner, for their invaluable help in making the study possible.

B. METHOD

1. *Stimulus Object*

The visual stimulus employed was a luminescent gold spot, 6cm. in diameter, which could be moved clockwise or counterclockwise in a horizontal plane to form an arc around *S*'s vertical axis. *S* was instructed to adjust this spot to his apparent median plane while he sat in a rotating chair. The apparatus was arranged so that the visual stimulus rotated with *S* and its position, relative to *S* during rotation, was remotely controlled by *E*.

2. *Remote Control*

The luminescent spot was attached at eye level to one end of a rod lying in the median sagittal plane of *S*. The other end of the rod was attached to the shaft of a commercially-built antenna rotor which revolved around a vertical axis parallel to that of *S*. The shaft of the rotor was connected by a chain drive to a selsyn actuator, which, in turn, acted in phase with a selsyn repeater at *E*'s desk. This selsyn system provided an indication of the position of the spot in half degrees. Thus, at the direction of *S, E* was able to manipulate the luminescent spot remotely while the chair was in motion and to read the position of the spot directly by means of the selsyn system.

The spot, together with the rod, rotor, and selsyn actuator, was placed on a platform attached to the arms of the rotating chair. The distance of the spot from *S* was 140 cm.; the distance of the rotor from *S,* 65 cm. The spot, therefore, moved in an arc of 75 cm. radius. Directly above the spot, a small luminescent blue cross provided a fixation point. The cross was permanently attached to the platform directly in front of *S* but, unlike the spot, could not be moved.

3. *Rotating Chair*

The rotating chair employed has been described in a previous publication (19). Only one speed, 25 rpm, was used for both clockwise and counterclockwise motion. The period of acceleration, from zero to a constant speed of 25 rpm, lasted 15 sec. The period of deceleration, from constant speed of 25 rpm to zero, lasted 9 sec. Attached to the chair was a head rest which fixed the position of *S*'s head so that a line from the corner of the eye to the ear canal was horizontal. Thus, the horizontal canals were approximately in a horizontal position during testing.

---

[2] For a review of the literature reporting this form of motion see Holsopple (12). Most recently Graybiel and his co-workers (9, 10) have investigated one form of such motion perception which they have termed the "oculo-gyral illusion." One aspect of their findings, the occasional report of displacement of a target light, bears directly on our investigation and will be considered in the discussion.

4. *Test Conditions*

All testing was carried out in the dark. Between trials, *S* was instructed to close his eyes and cover them with a pair of opaque goggles. During each trial, when *S*'s eyes were open, only the luminescent fixation cross and luminescent spot were visible.

Five test conditions were used: (*a*) control—absence of rotation (*C*); (*b*) acceleration in a clockwise direction (*A-CW*); (*c*) deceleration from constant clockwise rotation (*D-CW*); (*d*) acceleration in a counterclockwise direction (*A-CCW*); (*e*) deceleration from constant counterclockwise rotation (*D-CCW*).

Eighteen graduate male students were used as *S*s. Of these, nine were tested in the following sequence: *C, A-CW, D-CW, A-CCW, D-CCW;* and the other nine in the sequence: *C, A-CCW, D-CCW, A-CW, D-CW*. Each *S* was retested after a time lapse of at least one day.

5. *Measures Employed*

The position in which the spot appeared straight-ahead to *S* was measured in reference to the objective median plane. All measures were taken in degrees of visual angle. Deviations of the spot from the objective median plane in a clockwise direction (to the right of *S*) were arbitrarily designated "+", while those in a counterclockwise direction (to the left of *S*) were designated "—".

C. RESULTS

The results of both the test and retest, evaluated by analysis of variance and *t*-tests, are summarized in Tables 1 and 2 respectively. For both the test and retest the *F* ratio for experimental conditions is significant beyond the .01 level of confidence. The means indicate that in *A-CW* and *D-CCW* the spot is adjusted to the left of the objective median plane and control. Under these two conditions of stimulation, the spot, objectively in the median plane, is perceived as displaced towards the right. Since the task of *S* is to adjust the spot to his straight-ahead, he adjusts it to the left of the objective median plane. From this we may infer that the position of the apparent median plane of *S* has shifted to the left. We can also see from the means that in *A-CCW* and *D-CW* the position of the apparent median plane is shifted to the right of control. The *t*-tests indicate that these shifts are significant for both the test and the retest. All other comparisons are also significant, with the exception of those between *A-CW* and *D-CCW*, and *D-CW* and *A-CCW* where the displacements are in the same direction.

TABLE 1
SIGNIFICANCE OF DIFFERENCE OF MEANS AMONG EXPERIMENTAL CONDITIONS: TEST

Source of Variance	*df*		Mean square	*F*
	F-tests			
Individuals	17		1.89	——
a. Sequence		1	0.90	<1.00
b. Individuals within Sequence		16	1.95	<1.00
Conditions	4		94.87	24.92**
Error	68		3.81	
Total	89			

t-tests

		Mean differences			
Conditions	Means	A-CW	D-CW	A-CCW	D-CCW
Control (C)	0	2.1**	2.6**	1.8**	2.7**
Acceleration CW (A-CW)	—2.1		4.7**	3.9**	0.6
Deceleration from CW Rotation (D-CW)	+2.6			0.8	5.3**
Acceleration CCW (A-CCW)	+1.8				4.5**
Deceleration from CCW Rotation (D-CCW)	—2.7				

*Significant at or better than the .05 level of confidence.
**Significant at or better than the .01 level of confidence.

TABLE 2
SIGNIFICANCE OF DIFFERENCE OF MEANS AMONG EXPERIMENTAL CONDITIONS: RETEST

Source of Variance	*df*		Mean square	*F*
	F-tests			
Individuals	17		2.81	——
a. Sequence		1	0.26	<1.00
b. Individuals within Sequence		16	2.96	<1.00
Conditions	4		139.43	25.37**
Error	68		5.50	
Total	89			

t-tests

		Mean differences			
Conditions	Means	A-CW	D-CW	A-CCW	D-CCW
Control (C)	—0.3	2.1**	3.2**	2.8**	2.9**
Acceleration CW (A-CW)	—2.4		5.3**	4.9**	0.8
Deceleration from CW Rotation (D-CW)	+2.9			0.4	6.1**
Acceleration CCW (A-CCW)	+2.5				5.7**
Deceleration from CCW Rotation (D-CCW)	—3.2				

**Significant at or better than the .01 level of confidence.

The uncorrected reliability coefficients are presented in Table 3. These correlations between the same conditions in the test and retest indicate the significant reliability of the individual's performance in a given condition on the test to the same condition on the retest. The intercorrelations between all possible pairs of conditions in the test, retest, and test-retest are presented in Table 4. The upper left quadrant presents the intercorrelations within the test; the lower right quadrant, those within the retest; the upper right quadrant, those between the test and retest. This intercorrelation matrix manifests the high stability of individual performance obtained between conditions in our experiment and indicates that there are stable individual differences in susceptibility to acceleration and deceleration. Correlations between those conditions which cause a shift in position of the apparent median plane to the same side (between *A-CW* and *D-CCW*; and *D-CW* and *A-CCW*) are positive. Correlations between those conditions which cause a shift in the position of the apparent median plane to opposite sides (between *A-CW* and *D-CW*; *A-CW* and *A-CCW*; *D-CW* and *D-CCW*; and *A-CCW* and *D-CCW*) are negative. Thus, we may conclude that those *S*s for whom the position of the apparent median plane is displaced most towards the right ("+" adjustment) in *A-CCW* and *D-CW* tend to have the position of their apparent median plane displaced most towards the left ("—" adjustment) in *A-CW* and *D-CCW*. Analogously, *S*s for whom the position of the apparent median plane is displaced least toward the right in *A-CCW* and *D-CW*, tend to have the position of their apparent median plane displaced least towards the left in *A-CW* and *D-CCW*.[3]

TABLE 3

TEST-RETEST RELIABILITY COEFFICIENTS

Conditions	*r*
Control (C)	.56*
Acceleration CW (A-CW)	.88**
Deceleration from CW Rotation (D-CW)	.85**
Acceleration CCW (A-CCW)	.71**
Deceleration from CCW Rotation (D-CCW)	.80**

*Significant at or better than the .05 level of confidence.
**Significant at or better than the .01 level of confidence.

---

[3] No real contradiction exists between these findings and the low *F* ratios for individuals within sequence presented in Tables 1 and 2. Due to assigning "+" values for apparent median plane displacements towards the right and "—" values for apparent median plane displacements towards the left, a comparison of differences between subjects must necessarily result in a low *F* ratio if in fact the correlations are high and positive in "similar effect" conditions and high and negative in "opposite effect" conditions.

TABLE 4
INTERCORRELATIONS AMONG TEST-RETEST CONDITIONS

Conditions	Test A-CW	 D-CW	 A-CCW	 D-CCW	 C	 A-CW	Retest D-CW	 A-CCW	 D-CCW
C Test	.11	.33	—.24	.00	.56*	.20	.21	—.10	.24
A-CW Test		—.31	—.56*	.78**	.31	.88**	—.43	—.51*	.66**
D-CW Test			.29	—.52*	.50*	—.30	.85**	.38	—.39
A-CCW Test				—.55*	—.13	—.65**	.52*	.71**	—.76**
D-CCW Test					—.04	.75**	—.67**	—.72**	.80**
C Retest						.38	.31	.02	.15
A-CW Retest							—.49*	—.56*	.77**
D-CW Retest								.68**	—.66**
A-CCW Retest									—.80**
D-CCW Retest									

*Significant at or better than the .05 level of confidence.
**Significant at or better than the .01 level of confidence.

It might be mentioned here that an experiment, parallel to that described above was conducted under conditions of good illumination (the experimental room was illuminated by two overhead 150 watt bulbs). Although quantitative readings were not taken in this second experiment, it was clear that similar but, as we might expect, smaller displacements of the position of the apparent median plane occurred.

D. Discussion

The main finding of this study is that the position of the apparent median plane is significantly affected by accelerative and decelerative rotation around *S*'s vertical axis. Specifically, the position of the apparent median plane is displaced opposite to the direction of acceleration. In other words, when acceleration is clockwise, the displacement is counterclockwise; when acceleration is counterclockwise, the displacement is clockwise. Deceleration from rotation in one direction functions in a fashion equivalent to acceleration in the opposite direction. These relations are analogous to those obtained where apparent verticality is being measured (19).

The physical displacements summarized above are inferred from the position to which the spot was adjusted so as to appear straight-ahead. Obviously, if the spot was adjusted away from the objective median plane then it must have initially appeared to be displaced. This perceptual displacement effect must not be confused with a paradoxical form of illusory motion which also occurs under conditions of labyrinthian stimulation. Our *S*s report that the spot once displaced is seen to be constantly moving in the same direction as the displacement without further change in its relative position to the body. Graybiel, in his original paper on the "oculo-gyral illusion," discusses both of these effects and attempts to explain them by recourse to the conception of retinal displacement (10). He believes these effects to be due to the tracking of the image across the retina during ocular nystagmus. His basic reasoning is as follows: In acceleration or deceleration, nystagmic eye movements occur. In acceleration, the direction of the slow component of the eye movement is opposite to the direction of the rotation. Thus, in *A-CW* and *D-CCW* the slow movement is to the left; in *A-CCW* and *D-CW*, it is to the right. If the stimulus object is objectively stationary in relation to the subject, then, he argues, there will be a tracking of the image across the eye during the slow phase of the ocular movement. Phenomenally, what the subject will perceive, therefore, is a movement of the object opposite to the direction of the eye movement.[4]

Any attempt to interpret our results in terms of such a peripheral retinal explanation would encounter great difficulties. First of all, Graybiel's explanation does not distinguish between the phenomenon of displacement and that of illusory movement. It is difficult to see how an explanation in terms of tracking can simultaneously consider the perceived change in the location of the spot and the illusion that the spot is moving without further translocation. Secondly, if displacement were caused by movement of the image of the spot across the eye during the slow phase of nystagmus, what we would expect to see in *A-CW*, for example, is a movement of the spot from center to the right as the eye moved left during the slow phase, followed by the disappearance of the spot at the right during the fast phase, and its reappearance at the center as the next slow phase began. This kind of back-and-forth movement is not seen. Rather, our *S*s report that the spot is displaced to the right and remains in a fixed position relative to the observer throughout the course of acceleration.[5] It is precisely this fact, that the spot is displaced to a fixed position, which allowed us to measure its displacement in the way we did. Moreover, the high test-retest reliability coefficients which we obtained indicate that the displacements repeatedly occurred to the same position for each *S* in a given experimental condition.

---

[4] A provocative finding reported by Graybiel and confirmed by us is that during

There are two other more basic reasons why Graybiel's explanation or,

---

acceleration and deceleration, after-images are seen to be displaced in a direction opposite to that of the "real" visual stimulus light. Thus, for example, in *A-CW*, the luminescent spot is seen displaced to the right, but an after-image is seen displaced to the left of the objective median plane. Graybiel (10) discusses this fact as a strong argument for his interpretation of the "oculo-gyral illusion." An extended discussion of the relationship between after-images, "real" images, and the position of the eye is beyond the scope of the present paper. That the phenomenal localization of a "real" image or an after-image is not invariably related to eye position can, however, be readily demonstrated. If an object is fixated and the eye moved voluntarily, the object is not seen to be displaced. However, if the eye is moved passively with the finger, displacement of the object is seen. The obverse occurs with after-images. If an after-image is induced and the eye moved voluntarily, the after-image appears displaced. If the eye is moved passively with the finger, however, the after-image remains stationary. For a discussion of the pertinent literature on some aspects of this problem, see Helmholtz (11), Mach (15), James (13), Grant (8), Karrer and Stevens (14), and most recently, Bruell (4).

[5] The other aspect of the "oculo-gyral illusion," the illusory paradoxical motion without translocation, would also appear difficult to interpret with such a concept of ocular nystagmus. The illusory motion perceived is smooth and not jerky as we might expect it would be if due to retinal tracking. An experiment reported by Christian (5) in which *S*'s were rotated at an acceleration below that for perception of rotation is very enlightening on this point. Utilizing as a stimulus object two points of light close enough together so that they could only be discriminated within the foveal region (thus any eye movement would result in a fusion of the two points), Christian found that the points were seen distinctly apart but moving in the manner which we have described. For a further discussion of this phenomenon of illusory motion under conditions of labyrinthian stimulation and an attempted explanation of it see Mann (16). We shall discuss this problem further, and Mann's explanation of it, in a forthcoming publication reporting the influence of visual configurational dynamics on the apparent speed of such paradoxical motion.

for that matter, any purely retinal interpretation of our results, appears to be inappropriate. The first of these reasons has to do with the fact that we are dealing with the problem of the median plane, i.e., the problem of space localization and not the comparatively simpler problem of motion perception. Thus, although movement such as Wertheimer's phi phenomenon can possibly be adequately explained by a theory of retinal functioning, the perception of the position of objects in space, probably cannot. Space localization is far more than a retinal affair; it is dependent, among other things, upon such factors as head and trunk position. For example, if one orients one's head, trunk, and eyes straight-ahead, objects projecting onto the fovea will be located in the apparent median plane. However, if the head is now moved to the right or left, an object which now also projects on to the fovea will no longer appear in the apparent median plane. For a theoretical interpretation of the relationship between body position and space localization, the reader is referred to the discussions of Goldstein (7) and Werner and Wapner (20).

The criterion of maximum generality of a theory provides the final reason for not attributing our results to eye movements. A purely retinal theory might account for displacements obtained under conditions where the eye is free to move but it cannot account for analogous displacements observed when the eye is immobilized.[6] It is even more difficult for this kind of theory to account for the fact that various characteristics of the visual stimulus attended such as its position in space, (18, 22), its horizontal extent (3, 18), and its vectorial dynamics (21) will also affect the perception of the median plane. The analogous effects resulting from such apparently diverse factors as those mentioned above and labyrinthian stimulation seems difficult to understand without reference to some basic organismic concept to relate them, such as body tonus. As we have indicated in the introduction, body tonus is affected by labyrinthian stimulation. Since it has been shown that such characteristics of a visual stimulus as those referred to above will also cause innervation of the eye turning muscles (4) and, therefore, affect body tonus (20), we may conclude that the concept of body tonus fulfills the generality requirement. We believe it to be most probable that a theoretical formulation which provides for the interrelationship of the perception of space and of body tonus will best explain the perceptual displacement of the median plane found under such diverse conditions of stimulation.

---

[6] We refer here to Mach's classic study where putty was placed around the eye so as to immobilize it (15) and to those observations made with patients having paralysis of the rectus muscles (1, 11).

E. Summary

Two experiments were carried out to demonstrate that labyrinthian stimulation, induced by acceleratory and deceleratory rotation of *S* around his vertical axis, influences the visual perception of the median plane. Under conditions of acceleration, the position of the apparent median plane was displaced in the direction of rotation. Acceleration in one direction and deceleration from rotation in the opposite direction were found to act in an equivalent fashion.

The theoretical implications of these results were discussed. A retinal theory was found to be inadequate and was rejected in favor of the interpretation that body tonus affects the visual perception of space.

References

1. Battersby, W. S., Kahn, R. L., Pollack, M. R., & Bender, M. B. Effects of visual, vestibular, and somatosensori-motor deficit on autokinetic perception. *J. Exp. Psychol.*, 1956, **52**, 398-410.
2. Best, C. H., & Taylor, N. B. The Physiological Basis of Medical Practice. (5th ed.) Baltimore: Williams & Wilkins, 1950.
3. Bruell, J. H., & Albee, G. W. Effect of asymmetrical retinal stimulation on the perception of the median plane. *Percept. & Mot. Skills*, 1955, **5**, 133-139.
4. ———. Notes toward a motor theory of visual egocentric localization. *Psychol. Rev.*, 1955, **62**, 391-400.
5. Christian, P. Ueber unbewusste vestibulariswirkung. *Zsch. f.d. ges. Neurol. u Psychiat.*, 1939, **165**, 214-219.
6. Dusser, De Barenne, J. G. The labyrinthine and postural mechanisms. In C. Murchison (*Ed.*), *Handbook of General Experimental Psychology.* Worcester: Clark Univ. Press, 1934.
7. Goldstein, K. The Organism. New York: American Book, 1939.
8. Grant, V. W. Psychological Optics. Chicago: Professional Press, 1938.
9. Graybiel, A., Clark, B., MacCorquodale, K., & Hupp, D. Rôle of vestibular nystagmus in the visual perception of a moving target in the dark. *Amer. J. Psychol.*, 1956, **59**, 259-266.
10. Graybiel, A., & Hupp, D. The oculo-gyral illusion: A form of apparent motion which may be observed following stimulation of the semicircular canals. *J. Aviat. Med.*, 1946, **17**, 3-27.
11. Helmholtz, H. Helmholtz's Treatise on Physiological Optics: Vol. 3. The Perception of Vision. Southall, J. P. C. (*Ed.*). Rochester: Optical Society of America, 1925.
12. Holsopple, J. C. Space and the non-auditory labyrinth. In C. Murchison (*Ed.*), *The Foundations of Experimental Psychology.* Worcester: Clark Univ. Press, 1929.
13. James, W. The Principles of Psychology, Vol. 2. New York: Holt, 1902.
14. Karrer, E., & Stevens, H. C. The response of negative after-images to passive motion of the eyeball and the bearing of these observations on the visual perception of motion. *Amer. J. Physiol.*, 1930, **94**, 611-614.
15. Mach, E. The Analysis of Sensation. Chicago: Open Court, 1914.
16. Mann, C. W., Guedry, F. E., & Ray, J. T. Post-rotational perception of apparent bodily rotation. *J. Exp. Psychol.*, 1951, **41**, 114-120.

17. SHERRINGTON, C. S. The Integrative Action of the Nervous System. New York: Scribner, 1906.

18. WAPNER, S., WERNER, H., BRUELL, J. H., & GOLDSTEIN, A. G. Experiments on sensory-tonic field theory of perception: VII. Effect of asymmetrical extent and starting positions of figures on the visual apparent median plane. *J. Exp. Psychol.*, 1953, **46**, 300-307.

19. WAPNER, S., WERNER, H., & MORANT, R. B. Experiments on sensory-tonic field theory of perception: III. Effect of body rotation on the visual perception of verticality. *J. Exp. Psychol.*, 1951, **42**, 351-357.

20. WERNER, H., & WAPNER, S. Toward a general theory of perception. *Psychol. Rev.*, 1952, **59**, 324-338.

21. ———. Studies in physiognomic perception: I. Effect of configurational dynamics and meaning induced sets on the position of the apparent median plane. *J. of Psychol.*, 1954, **38**, 51-65.

22. WERNER, H., WAPNER, S., & BRUELL, J. H. Experiments on sensory-tonic field theory of perception: VI. Effect of the position of head, eyes, and of object on the position of the apparent median plane. *J. Exp. Psychol.*, 1953, **46**, 293-299.

Department of Psychology
Brandeis University
Waltham 54, Massachusetts

Published as a separate and in *The Journal of Psychology*, 1959, **47**, 37-40.

SEXUAL RÔLE, THE SHORT STORY, AND THE WRITER*

Department of Psychology, University of California, Santa Barbara

WILLIAM D. ALTUS

A. PROBLEM AND PROCEDURE

Reading articles and short stories in the *New Yorker* magazine is something of a pig-in-a-poke affair—that is, the name of the author comes out in a discreet murmur at the end rather than in a breezy shout at the very beginning which is typical of the conventional magazine. Years ago the writer came to the idle and quite possibly egotistical conviction that he could frequently distinguish the sex of the author of *New Yorker* short stories before he had read more than the first few sentences. Such an assumption carries the implication, of course, that masculinity-femininity is a pervasive affair, so pervasive that it would color indelibly even the writings of the professional whom one would expect to be on guard. If such an assumption could be shown to be a reasonable inference from objective data, then one would be constrained to grant that the psycho-social sexual rôle imposed on us all through cultural processes is strong indeed. In the present study that assumption becomes the hypothesis—namely, that there is something in the manner in which writers compose the opening sentences of their short stories which gives their sex away to a degree beyond what mere chance would allow.

For a preliminary tryout of this hypothesis, it was decided to choose 200 short-story openings at random from two sources, the *New Yorker* and the *Saturday Evening Post*. The size of the excerpt was limited to a maximum of 80 words; if the opening sentences were short, no more than three sentences would be quoted, even though only 15 or 20 words might be involved. The 100 selections from the *New Yorker* came from the period 1946-1950; the 100 from the *Post* from 1951-1954. The rationale for going back somewhat in time was to avoid the possible adulterating effects of the reader's memory for the recently perused short story. Bias in selecting the excerpts was avoided through using a paid assistant who randomized the choice of authors within the limits imposed by keeping the writers evenly balanced as to sex and by the stricture that no author was to be used more than once.

---

*Received in the Editorial Office on September 24, 1958, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

The *Post* was chosen as a supplementary source since in a number of obvious ways it is markedly divergent from the *New Yorker*: Circulation, type of audience, and so it is rumored, a "slicker" type of story.

B. Results

In the preliminary tryout of the 200 items, 100 college undergraduates participated. A quartile analysis followed, with all four quartiles analyzed, the criterion for the analysis being the number of authors correctly inferred. Eighty-six of the 100, incidentally, earned a better-than-chance score—that is, correctly inferred more than 100 of the 200 authors. In the quartile analysis, only those selections which showed a consistent increment from the lowest to the highest one-fourth were considered for retention. This was done to obviate as well as could be the possibility of chance causing the retention of certain items; when items are essentially of the true-false type, as they are in this instance, the inclusion of the middle quartiles in an analysis helps control this adventitious factor. The object in selecting items for retention was to discriminate between those who did well at this type of inference versus those who did not do so well. In this way, the differences between high and low scores should be maximized.

Thirty excerpts finally remained, 15 with male authors, 15 with female. Some 265 college students, 187 women and 78 men, were cozened into taking the new and much shortened form. At the same time they were given a difficult, college-level, multiple-choice vocabulary test as a check on the relation of verbal ability to the ability to infer the sex of a short-story author.

Roughly four out of five students (78.9 per cent) in the new population earned better than a chance score on the new 30-item inference test. Eighty-four per cent of the women correctly inferred 16 or more of the 30 authors—that is, better than a chance score; only 66.7 per cent of the men were so intuitive, if, indeed, intuitive is the correct term to use here. As might be expected the significance of the mean differences between male and female students for correctly inferring the author's sex reaches statistical significance: the t-value is 5.36. Women apparently do better at smelling out the true sex of an author when only a snippet of his professional writing is present. This superiority is not due, however, to superior verbal intelligence, since the males earned a significantly higher (.05 l.c.) mean vocabulary score than did the larger group of college females who participated. There was only a slight relation between vocabulary score and the number of correctly inferred sex choices: For the men the r (PPM) was .26 between the 20-item, multiple-response vocabulary test and the number of correct author-sex inferences; the women, the *r* was only .14; for the total group of both sexes the *r* was .12. Oddly enough, however, the male superiority in verbal intelligence was not related to superiority in shrewdly inferring an author's sex from exposure to a few of his words.

Quizzing the high and lower scorers on the inference test and examining the 30 items which were winnowed out of the original matrix of 200 for possible clues as to the rationale of success in the type of task here considered led to the conclusion that probably three cues were significant in differentiating the successful from the unsuccessful: the style of the selection, the type of thing or action depicted (content), and, finally, where it was patent, the sex of the described character. For instance, preciosity of phrasing and colorful adjectives were used in some instances as a clue to femininity, apparently. If the type of thing described was sex-bound, the content might be revealing. In one of the original 200 items, there is a description of salmon fishing; practically everyone guessed correctly it was a male author. Finally, there seems to be a tendency (not quantified as yet) for female writers to describe females in their opening sentences and vice versa. This cue, usually employed without conscious awareness, was the basis for a part of the relatively high scores earned by some.

C. DISCUSSION

The data give basis for a belief that the masculinity-femininity variable, as was hypothesized, seeps through even in the first few words of a professional short-story writer, whether he writes for the relatively sophisticated *New Yorker* or a mass circulation magazine such as the *Saturday Evening Post*. Seventeen of the 30 excerpts chosen from the original 200 came from the former magazine, 13 from the *Post*. The type of magazine, therefore, appears to be of little import in the present research.

It appears likely that women, at least college women, do somewhat better at such kinds of literary inference than do men. Such superiority may possibly derive from the greater interest on the part of women in things verbal in contrast to the quantitative. One seldom finds a female physics or mathematics major; sexual stereotyping is markedly effective in such curricular choices. Female English majors abound. Perhaps the college woman has had more short-story reading in her background. It may be, too, that women are more sensitive to subtle verbal nuances, again through differential training imposed by the psycho-social sexual rôle. The woman who is successful in sniffing out the right answer from minimal cues, verbal or otherwise, in interpersonal relations is the same person, one may safely assume, who is successful as an individual in getting what she wants in her socially-defined rôle as a woman.

Verbal intelligence, at least in a relatively homogeneous college population, seems linked but slightly to the ability to infer the sex of an author. The men, who were significantly poorer at such inference, actually had a superior verbal ability, as measured by a vocabulary test. Within sex groups, however, there was a slight tendency for those with the higher vocabulary scores to do somewhat better in matching the literary excerpt with the correct sex of the author.

The college students who were most successful inferrers of male-female authorship of short stories were generally unable, under questioning, to explain precisely the factors involved in their success. To them it was primarily an "art," despite their frequently lengthy disquisitions when pressed for an explanation. Close study of the evidence presently available constrains one to hypothesize that at least three kinds of cues may have been operating: the style of the writer, the content of the excerpt, the sex of the principal person described. This part of the problem needs further investigation.

As is usual in such matters, more questions than answers are turned up. One would like to know whether the person who is quite effective at inferring an author's sex when only minimal cues are present is effective through the realm of the traditionally intellectual or through what is commonly thought of as personality factors. One would like to know, also, whether wide and diverse reading is associated with such an inferring ability. It seems reasonable to assume that the authorship of themes written by the college student who is much younger and less sophisticated than the professional short-story writer should be easier to infer; but research is lacking to substantiate such a hypothesis. Intriguing, also, is the question as to why certain authors are easier to spot as male or female and whether such typing is due to the author's greater acculturation in his socially defined rôle. These are all questions to which, as yet, the only answers are hypothesis and conjecture.

Department of Psychology
University of California, Santa Barbara
Goleta, California

Published as a separate and in *The Journal of Psychology*, 1959, **47**, 41-48.

THE VERTICAL-HORIZONTAL ILLUSION IN ARTIFICIAL VISUAL FIELDS*

Department of Psychology, University of Stockholm

THEODOR M. KÜNNAPAS[1]

A. PROBLEM

The overestimation of the vertical direction was studied in a series of investigations (2, 4, 5, 6, 7). We found that this overestimation is due to the oval shape of the visual field which is extended in a horizontal direction. So the shape of our visual field is one of the factors which influences the subjective estimation of the vertical direction.

This our general hypothesis has been tested and confirmed in cases of the normal visual field. A further way to test it would be to investigate artificial fields.

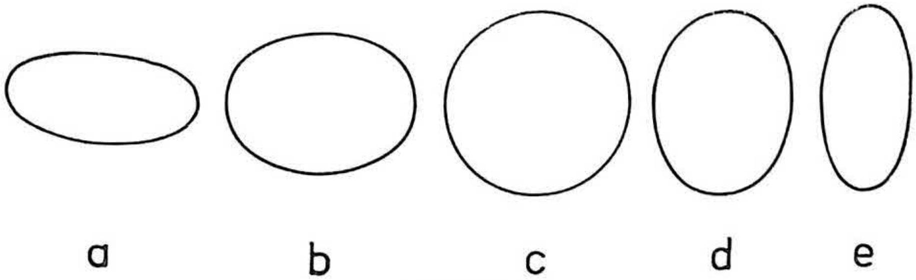


FIGURE 1
DIFFERENT SHAPES OF THE ARTIFICIAL FIELDS
(*a*) The most elliptical lying oval with the ratio V/H equal to 0.50. (*b*) The less elliptical lying oval with the ratio V/H equal to 0.75. (*c*) The circle-shaped field where the ratio V/H is 1.00. (*d*) The less elliptical standing oval with the ratio V/H equal to 1.33. (*e*) The most elliptical standing oval where the ratio V/H is 2.00.

The normal binocular visual field has the shape of a lying oval where the ratio of the vertical axis to the horizontal (V/H) is approximately 3 : 4 or 0.75. A systematic variation of this ratio must influence the size of the overestimation of the vertical line (OV), according to our hypothesis. With the increase of the ratio it should be expected that the OV decreases, and vice versa (3). Figure 1 shows five different artificial visual fields. The oval (a) has the lowest V/H. In the following fields the ratio increases

---

*Received in the Editorial Office on September 27, 1958, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This investigation was supported by a grant from Magnus Bergvall's Foundation.

successively so that the last oval (e) has the highest V/H. According to our hypothesis, it is to be expected that the OV is, on the contrary, successively decreasing from the oval (a) to the (e). These five visual areas are chosen so that one field (b) has approximately the same shape as the normal visual field. So it is to be expected that in oval (b) the OV would be the same as in the normal field. In the oval (a) the OV should be greater and in the others smaller as compared with the normal.

The purpose of the present study is to investigate quantitatively how the artificial visual fields with different ratios of V/H influence the OV, and thereby to provide a further test of our general hypothesis.

B. METHOD

1. *Apparatus*

An *L*-shaped vertical-horizontal figure was exposed in the central part of a round white surface with a radius of 300 mm. The horizontal line was the constant stimulus and it was 50 mm long. The length of the vertical line could be adjusted by the subject. Both lines were 0.5 mm wide. The figure was visible only when the light was switched on by the experimenter. Details have been given in previous papers (3, 4, 7).

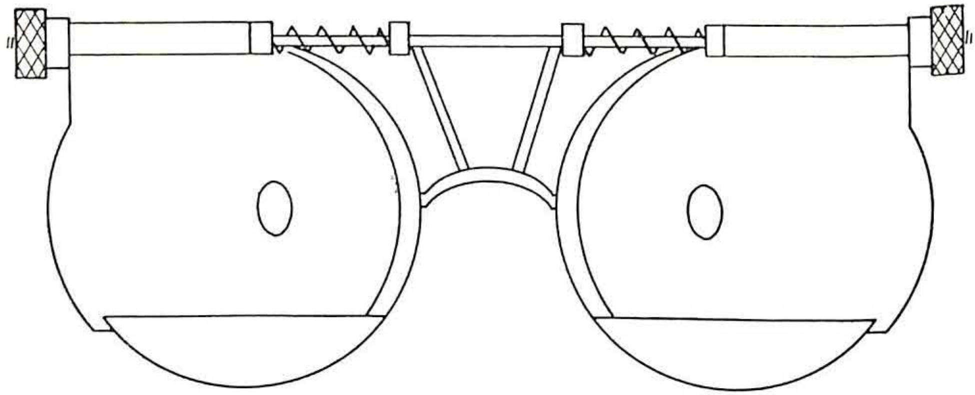


FIGURE 2

THE SPECTACLES BY WHICH THE ARTIFICIAL VISUAL FIELDS WERE PRODUCED

In the illustration the openings in the spectacles have the shape (d) with the less elliptical standing oval.

The normal binocular visual field was artificially limited by special spectacles. A schematic front-view of the spectacles is given in Figure 2. The glasses were replaced by aluminium shields. Each of both shields had the same opening which was nearer to the temporal boundary. Five different pairs of openings were used in the experiments: two lying ovals, one circle, and two standing ovals (Figure 1). The vertical and horizontal axes of the lying ovals were: (*a*) 2 mm and 4 mm, and (*b*) 3 mm and 4 mm. The diameter of the circle (*c*) was 4 mm. The axes of the standing ovals were: (*d*) 4 mm and 3 mm, and (*e*) 4 mm and 2 mm. So the shape of the ovals (a) and (e) was the same, and the latter was only rotated 90°. Also the oval (d) had the same shape as (b), rotated 90°. The V/H of the openings was: (*a*) 0.50, (*b*) 0.75, (*c*) 1.00, (*d*) 1.33 and (*e*) 2.00. The shields were used in pairs, each member of the pair containing the same shape of opening. The distance between the openings was adjustable to the pupillary distance of the subject, and also to the distance from the eyes to the test figure, so that the two openings of the left and of the right eye gave one field. The spectacles were attached to the head by an elastic band. Both eyes were covered sidelong by a device of the spectacles so that the *S* could see only through two openings. The weight of the spectacles was 92 g.

2. *Procedure*

The subject wearing spectacles sat in front of the apparatus in a fixed position and his head was held steady by a chin rest. The distance from the *S*'s eyes to the test figure was 600 mm. The figure itself was at eye level so that the point where the horizontal and vertical lines intersect was exactly between the *S*'s eyes.

The method of adjustment was used. The *S* was instructed to view the figure so that the intersection point of the two lines was always in the centre of the artificial visual field. He had to compare both lines of the exposed vertical-horizontal figure with each other, and to adjust the length of the vertical line so that it appeared subjectively equal to the horizontal line. The *S* saw the figure only when he was adjusting the vertical line. After each adjustment the figure was switched off and rearranged by the *E* before the next exposure.

The *S* viewed the figure with both eyes. In preliminary trials *S*s were made familiar with the experimental situation.

3. *Experimental Conditions*

Six conditions were arranged:

a. A-Condition. In this condition the artificial visual field had the shape of the most elliptical lying oval (a) where the ratio V/H of the axes was 0.50 (Figure 1, a). On the white surface of the apparatus the vertical axis of the artificial visual field was approximately 215 mm long and the horizontal axis 430 mm.

b. B-Condition. Here the artificial field had the shape of the lying oval (b) which was less elliptical (Figure 1, b). The ratio V/H was 0.75. The vertical axis of the field was approximately 323 mm and the horizontal axis 430 mm.

c. C-Condition. The circle-shaped artificial visual field with the ratio V/H equal to 1.00 (Figure 1, c) had the diameter 430 mm on the white surface of the apparatus.

d. D-Condition. Here the form of the artificial field was the standing oval (d) where the ratio V/H was 1.33 (Figure 1, d). The visual field of this condition had the same shape and size as in B-Condition, only rotated 90°. So the vertical axis was longer than horizontal. On the surface of the apparatus the former was 430 mm long and the latter 323 mm.

e. E-Condition. In this condition the visual field had the shape of the standing oval (e) with the ratio V/H = 2.00 (Figure 1, e), which was the same as in A-Condition, only rotated 90°. The vertical axis was approximately 430 mm long and the horizontal axis 215 mm.

f. N-Condition. In this condition *S*s had the normal binocular visual field, without the artificial spectacles. The ratio V/H was here approximately the same in the B-Condition. Contrary to the B-Condition the visual field was not limited.

The sequence of the conditions was N, B, E, C, A, D—D, A, C, E, B, N. All *S*s were given the same sequence.

4. *Subjects*

Six men and 10 women, most of them university students of psychology, took part in the experiments. All *S*s had normal visual acuity. Subject No. 7 had a weak astigmatism. Each *S* made 16 adjustments with each of six conditions: 8 ascending and 8 descending. Thus each *S* made 96 adjustments and the whole group made a total of 1,536 adjustments.

C. RESULTS AND DISCUSSION

The results for each *S* and for each condition are shown in Table 1. The data represent the average length of the vertical line (in mm) which appeared equal to the horizontal line of 50 mm for each of the six experimental conditions. The adjusted length of the vertical line is different in different conditions. It is shortest in A-Condition and increases successively from B- to E-Condition. In the latter condition it is longest. As compared with the normal visual field (N), the adjusted length of the vertical line in the artificial fields is shorter in (A), nearly the same in (B), and successively greater in others, from (C) to (E). The results of a few *S*s, however, more or less deviate from this general tendency.

The size of the OV is given in Table 2. The illusory phenomenon OV is greatest in A-Condition and successively decreases from B- to E-Condition where it is smallest. The OV of the B-Condition is nearly the same as in the normal visual field (N). This result was also to be expected. The artificial visual field of the B-Condition was chosen approximately with the same ratio V/H of the vertical and horizontal axes as that of the normal visual field. The artificial limitation of the normal field to a smaller area was the only difference between these two conditions. So the experimental data confirmed that the spectacles could not in any way distort the results.

TABLE 1

LENGTH OF THE VERTICAL LINE IN MILLIMETERS APPEARING EQUAL TO A HORIZONTAL LINE OF 50 MM IN THE NORMAL AND IN DIFFERENT ARTIFICIAL VISUAL FIELDS

Subject	Normal (N)	Ratio V/H of the axes				
		(A) 0.50	(B) 0.75	(C) 1.00	(D) 1.33	(E) 2.00
1	49.09	48.94	48.78	49.31	50.13	49.78
2	48.56	47.44	47.63	47.88	47.84	48.03
3	45.91	44.94	46.09	45.97	46.03	46.63
4	46.97	47.50	47.53	47.53	47.66	48.09
5	49.28	49.31	50.00	49.94	50.13	50.91
6	48.44	45.72	47.50	47.03	47.59	47.88
7	48.13	47.41	47.56	47.84	47.78	48.41
8	48.34	47.34	47.94	47.75	47.88	48.25
9	48.00	47.44	48.03	48.69	48.16	48.72
10	47.09	46.19	46.38	46.72	46.50	46.91
11	47.91	48.53	49.84	50.16	50.56	51.25
12	48.31	47.88	48.38	48.28	48.50	48.69
13	48.66	48.50	48.53	48.28	48.25	48.31
14	47.53	47.03	48.06	47.94	48.91	48.28
15	48.38	46.78	47.81	48.75	50.63	52.75
16	47.72	46.47	47.00	48.25	48.34	49.75
Mean	48.02	47.34	47.94	48.15	48.43	48.92

TABLE 2

OVERESTIMATION OF THE VERTICAL LINE (OV) IN THE NORMAL AND IN DIFFERENT ARTIFICIAL VISUAL FIELDS

Visual field	OV	
	mm	Per cent
Normal (N)	1.98	4.0
Artificial (A)	2.66	5.3
Artificial (B)	2.06	4.1
Artificial (C)	1.85	3.7
Artificial (D)	1.57	3.1
Artificial (E)	1.08	2.2

All other artificial visual fields (A), (C), (D) and (E) had different shapes from the normal. So one may expect that the OV in these conditions is different from that of the normal field. This is confirmed by the data in Table 2. In one case, in the A-Condition, where the ratio V/H of the axes is smaller than in the normal, the illusion is also greater. In other cases, where the ratio V/H is greater than in the normal the OV is accordingly smaller.

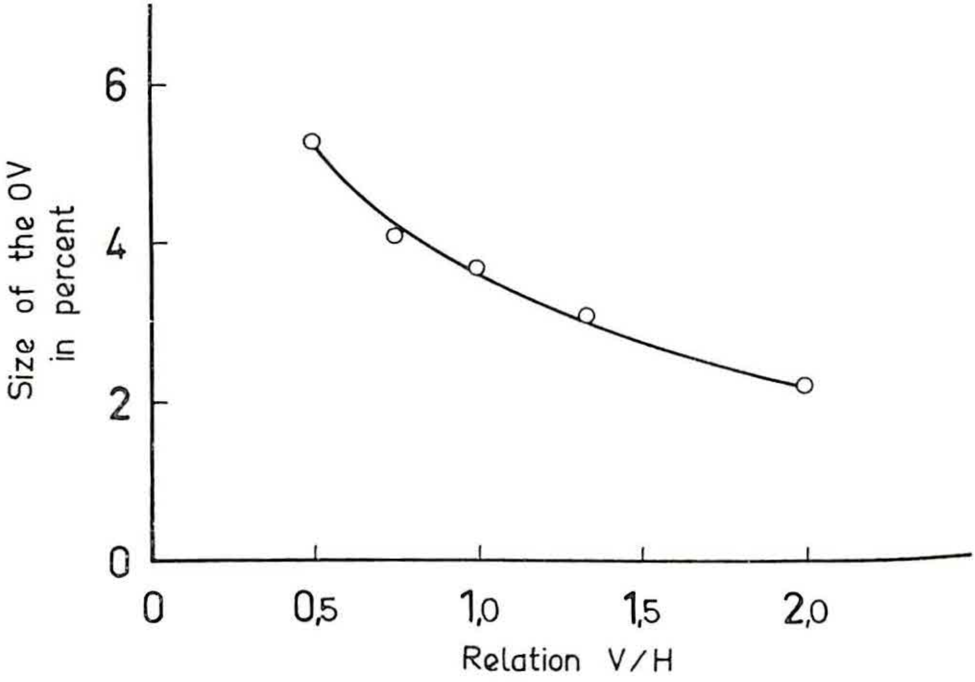


FIGURE 3
THE OVERESTIMATION OF THE VERTICAL LINE (OV) AS A FUNCTION OF THE RATIO V/H OF THE VERTICAL AXIS TO THE HORIZONTAL

In Figure 3 the size of the OV in per cent is plotted against the ratio V/H. The curve shows that the size of the OV is a function of the ratio V/H of the two axes.

The size of the differences and their statistical significance is summarized in Table 3. There are different levels of significance. Both most elliptical artificial fields produce differences which have a high significance, especially the difference between most and less elliptical lying fields and the difference between circle and most elliptical lying oval. It seems that it is as easy to increase as to decrease the OV. The differences between the circle and both less elliptical fields are not significant. The difference between normal and less elliptical field (b) has an especially low *t* value.

TABLE 3

SIZE (IN MILLIMETERS) AND SIGNIFICANCE OF THE DIFFERENCES ($df = 15$)

Difference	mm	*t*
(E) and (D)	.49	3.06**
(D) and (C)	.28	2.00
(C) and (B)	.21	1.75
(B) and (A)	.60	4.44***
(E) and (A)	1.58	4.39***
(D) and (B)	.49	2.45*
(E) and (C)	.77	3.21**
(C) and (A)	.81	5.06***
(N) and (B)	.08	.35

***$P \leqslant .001$
**$P \leqslant .01$
*$P \leqslant .05$

In the conditions (D) and (E) the horizontal axis of the artificial visual field is shorter than the vertical axis and one may expect that the OV changes into an overestimation of the horizontal line (OH), at least in the (E), where the horizontal axis is shortest. Although the OV decreases in (D) and more in (E), however, it does not change into an OH. This discrepancy may be explained by two assumptions. (*a*) Habituation to the OV is operating in the opposite direction and is preventing the change of the OV into the OH. (*b*) There are some other factors which influence the OV and which do not vary with the variation of the V/H in the artificial fields, and therefore prevent the OV changing into an OH. A previous paper (7) points in the latter direction. From the fact that when the head is tilted horizontally the OV changes into an OH, one may assume that there is a factor other than the oval shape of our visual field which influences the OV. This factor is possibly to be found in the eye itself.

We may conclude that our general hypothesis is supported by the present experiments. The oval shape of the visual field, however, is not the only factor which influences the OV.

D. SUMMARY

A further attempt has been made to provide a partial test of our general hypothesis according to which the overestimation of the vertical direction is due to the oval form of our visual field with its longer horizontal axes.

Our normal binocular visual field was compared with different artificial visual fields which were produced by specially constructed spectacles. Five

artificial fields were used where the ratio of the vertical axis to the horizontal (V/H) was 0.50, 0.75, 1.00, 1.33 and 2.00.

It was found that the overestimation of the vertical line (OV) is a function of the V/H. The OV decreases with an increase of V/H. It was also found that none of the artificial fields could change the OV into an overestimation of the horizontal line (OH), not even the field with the largest V/H.

It was concluded that OV is influenced by the oval shape of our visual field, but that this factor is not the only one.

REFERENCES

1. GIBSON, J. J. The Perception of the Visual World. Boston: Houghton-Mifflin, 1950.
2. KÜNNAPAS, T. M. An analysis of the "vertical-horizontal illusion." *J. Exp. Psychol.*, 1955, **49**, 134-140.
3. ———. Influence of frame size on apparent length of a line. *J. Exp. Psychol.*, 1955, **50**, 168-170.
4. ———. The vertical-horizontal illusion and the visual field. *J. Exp. Psychol.*, 1957, **53**, 405-407.
5. ———. Vertical-horizontal illusion and surrounding field. *Acta Psychol.*, 1957, **13**, 35-42.
6. ———. Interocular differences in the vertical-horizontal illusion. *Acta Psychol.*, 1957, **13**, 253-259.
7. ———. Influence of head inclination on the vertical-horizontal illusion. *J. of Psychol.*, 1958, **46**, 179-185.

Psychological Institute
University of Stockholm
Teknologgatan 8
Stockholm Va, Sweden

Published as a separate and in *The Journal of Psychology*, 1959, **47**, 49-56.

SOME PERCEIVED PROPERTIES OF THE DIFFERENCE BETWEEN LEADERS AND NON-LEADERS*

Department of Psychology, Queens College, New York

MICHAEL BEER,[1] ROBERT BUCKHOUT,[1] MILTON W. HOROWITZ, AND SEYMOUR LEVY[2]

A. INTRODUCTION

The functional concept of leadership stresses that leadership is delegated by the group and is conferred upon that individual who, under the present conditions of the group's growth and development, is best able to meet its needs. Essentially, this means that leadership will vary as the group's task or conditions vary and those attributes which may be perceived as ideal leadership qualities under one condition may, under other conditions, fail to satisfy the group's needs.

Nevertheless, it is true that many ongoing groups have achieved relative structural and internal psychological stability in which the characteristics of perceived leadership qualities remain relatively constant. It was the purpose of this study to explore experimentally some of these attributes of leadership. A review of the literature and a logical analysis of the differences between leaders and non-leaders lead to those general personal attributes which are seen as desirable qualities of leaders under a fairly wide range of stable psychological conditions. Three areas which seemed most relevant to the experimenters were self acceptance, need achievement, and interpersonal skills.[3]

1. *Self Acceptance*

Self acceptance includes three sub-categories. These are: (*a*) a feeling of personal worth, (*b*) self confidence, (*c*) a realistic outlook (toward self and others).[4]

The self accepting person sees himself as generally positive but also, and more importantly, sees relatively few, if any, negative areas in his self-image.

---

*Received in the Editorial Office on October 2, 1958, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] Now at North Carolina State College.

[2] Now at Pittsburg Mills, Minneapolis, Minnesota.

[3] Although these three categories seem to be logically independent, items designed to tap these categories are not so easily separated. It may be noted, too, that interpersonal skills, to an extent, is an expression of the other two categories.

[4] The realistic outlook was measured not only by the rating scale items, but also by the difference between the mean self ratings of leaders and non-leaders with the mean group ratings on the entire 26 items.

He is confident, with a background of success experiences and, finally, he has come to terms with life to the extent that he sees himself objectively and is generally realistic about his environment. For such a person, criticism is not seen as a threat to the self. Having a high degree of personal worth he is not overly dependent upon affection from others, and hence, has relative independence of action. Being more secure, he has a more realistic outlook and is better able to make decisions without being unduly swayed.

2. *Need Achievement*

This category is clearly motivational, subsuming factors like drive and ambition but it includes cognitive factors, too. A person with high need achievement has relatively high clarity of path and goal and is perceived by the group as having the ability to direct it to its goals. Further, a person with high need achievement accepts responsibilities and perhaps even seeks out new ones.

3. *Interpersonal Skills*

This area includes four sub-categories. These are: (*a*) warmth, (*b*) outgoingness, (*c*) tact, (*d*) dominance.

In part, interpersonal skills represent the mechanics of leadership. To some extent, they are the tools whereby self acceptance and need achievement come to fruition. Part of perceived leadership is not only the ability to make people feel at ease (warmth) but also to be diplomatic (tact) when the occasion demands it. It was felt that such a person would be extraverted and active and be able to start conversations easily (outgoingness). He would also be persuasive and have the ability to influence others (dominance).

B. PROCEDURE

Twenty-six items which were designed to tap the three major areas and their sub-categories were assembled into an instrument shown below. Of the three major areas, self acceptance is represented by the first nine items equally distributed among the three sub-categories. Need achievement is represented by Items 10-14. The next 12 items represent interpersonal skills with the four sub-categories containing three items each. The rating scales (Table 1) are each 90 mm. in length with two anchors and a midpoint and follow the procedure suggested by Champney and others (1, 2).

The approach used in this study was to compare a leader and a non-leader in the same group. The sample of compared persons consisted of 10 elected leaders of social groups (fraternities, sororities, and male and female houseplans) and 10 members of the same groups who had never held an elected

TABLE 1

Place a check mark (√) any place along the line that best expresses your feeling about the person you are rating. Check all scales. Do not spend too much time over any scale. Let youı initial feelings decide where the check mark should go.

1.	easily embarrassed	——|——————|——————|——	takes everything in stride
2.	wants everyone's love	——|——————|——————|——	independent
3.	touchy, easily hurt	——|——————|——————|——	able to take criticism
4.	self confident	——|——————|——————|——	lacks self confidence
5.	leans on people; easily lead	——|——————|——————|——	self reliant
6.	faith in own ability	——|——————|——————|——	unsure of self
7.	uses good judgment	——|——————|——————|——	uses poor judgment
8.	ideas ar usually unrealistic	——|——————|——————|——	ideas are usually realistic
9.	judgments are usually incorrect	——|——————|——————|——	judgments are usually correct
10.	looks ahead	——|——————|——————|——	not concerned about the future
11.	accepts responsibility reluctantly	——|——————|——————|——	accepts responsibility easily
12.	tolerates frustration	——|——————|——————|——	can't tolerate frustration
13.	goal is clear	——|——————|——————|——	goal is unclear
14.	easy going	——|——————|——————|——	drives self
15.	makes you feel at ease	——|——————|——————|——	makes you feel uncomfortable
16.	cold	——|——————|——————|——	warm
17.	big hearted	——|——————|——————|——	selfish
18.	often starts conversation	——|——————|——————|——	rarely starts conversation
19.	active	——|——————|——————|——	passive
20.	unobtrusive	——|——————|——————|——	stands out
21.	diplomatic	——|——————|——————|——	undiplomatic
22.	impulsive	——|——————|——————|——	controlled
23.	inconsiderate	——|——————|——————|——	considerate
24.	persuasive	——|——————|——————|——	unpersuasive
25.	easily influences	——|——————|——————|——	is easily influenced
26.	forceful	——|——————|——————|——	nonassertive

office in any known situation and who were equated for semester in college. Five of the groups were male and five were female. Being an office holder is not necessarily a good criterion of leadership. It is a pragmatic one, however, and does assure us that at the time these data were obtained, those who were designated as leaders were probably giving their groups the best solution to their needs at that time.

The data were obtained during the regularly scheduled meetings of the social groups. Ten members of the group (randomly selected except that they included the leader and the non-leader) were given a copy of the instrument and were instructed to rate the person whose name appeared on the sheet on all 26 items. This was done twice. The first time the scales had the name of the leader on them and the second time the scales had the name of the non-leader on them. Hence, ratings were obtained by the group on the leader and the non-leader. In addition, of course, the leader's and the non-leader's ratings constituted self ratings.

The subjects were instructed that if their own name appeared on the rating sheet to rate themselves and to make a check mark. This prevented them from knowing that they were all rating the same person and tended to avoid undue speculation about the meaning of the procedure at the time of testing. In addition, the scales were handed out in such a way that half of the group rated the leader first and half rated the non-leader first. The self-ratings of the leader and the non-leader were used only as a measure of how reality oriented these two persons were in comparisons of their self ratings with the group ratings.

C. RESULTS

The means for each of the 26 items were computed for the group ratings of the non-leader and leader. In addition, a "*D*" score (the difference between self ratings of leaders and non-leaders, and mean group ratings) was computed for each item and for the summed scores of the 26 items. Statistical significance of differences were determined by *t*-tests.

Twenty-three of the 26 items on the scale gave differences in the expected direction, the leaders being rated higher than the non-leaders. Fourteen of these differences, of which two were reversals, were significant at or beyond the .05 level (Table 2).

In the self acceptance category five of the items (out of nine) showed significant differences (see Table 2). The data show that the degree of confidence, and the degree of realism as rated by the group is higher for leaders than for non-leaders. Item 6 (Faith in own ability) gave the largest

TABLE 2
SELF ACCEPTANCE ITEMS

Item No.	Statement	Mean ratings Leader	Mean ratings Non-leader	p
1.	Takes everything in stride	48.7	56.6	$<.05$
5.	Self Reliant	62.2	53.2	$<.05$
6.	Faith in own ability	65.4	51.4	$<.01$
7.	Uses good judgment	75.9	55.9	$<.05$
8.	Ideas usually realistic	61.0	52.7	$<.01$

difference between leaders and non-leaders. The items designed to measure personal worth were, in general, the poorest discriminators. For example, Item 1 yielded a reversal which was significant. This might be explained by the fact that the leader is constantly before the group and therefore is more frequently put in situations where there is a possibility for embarrassment. In addition, of course, the non-leader retains anonymity whereas the leader is visible. Two of the three items designed to tap the realistic outlook were significant. The "*D*" score that was obtained and which is discussed below further substantiates that this is an important factor in differentiating leaders and non-leaders.

In the area of need achievement (Table 3), three out of five items were significant. However, one item (goal clarity) was reversed, non-leaders being rated as having greater goal clarity than leaders. There is no obvious explanation for this. The other two items (11 and 14) indicate that leader's tend to accept responsibility more easily and have more drive than non-leaders. Accepting responsibility easily and drives self are terms that indicate willingness to help, eagerness and ambition.

TABLE 3
NEED ACHIEVEMENT ITEMS

Item No.	Statement	Mean ratings Leader	Mean ratings Non-leader	p
11.	Accepts responsibility easily	64.7	48.4	$<.01$
13.	Goal is clear	44.8	53.6	$<.05$
14.	Drives self	37.3	28.2	$<.01$

All but one of the scales in the area of Interpersonal Skills (Table 4) yielded results in the expected direction, and six of the 12 items showed statistically significant differences. In Items 24, 25, and 26 (Dominance), in particular, the leaders were rated significantly higher than the non-leaders on their persuasiveness, their ability to influence others, and their forcefulness. There were no statistical significant items in the sub area warmth (Items 15, 16, 17). It is interesting, however, that there was a

TABLE 4
INTERPERSONAL SKILLS ITEMS

Item No.	Statement	Group rating Leader	Non-leader	p
19.	Active	68.1	62.4	$<.05$
21.	Diplomatic	64.0	52.5	$<.01$
22.	Controlled	54.1	46.6	$<.05$
24.	Persuasive	56.0	49.3	$<.05$
25.	Easily influenced	52.6	45.0	$<.01$
26.	Forceful	54.0	47.5	$<.05$

non-significant reversal in that non-leaders were rated as being "warmer" than leaders, with the largest difference occurring in the female groups studied. On Items 21 and 22 (tact), the leaders were rated as being more diplomatic and better controlled than the non-leaders. Leaders were also rated as being more active. It is apparent that these interpersonal skills are highly valued by the groups used in this study and probably are large determiners of election to leadership.

In comparing the self ratings by the leaders and non-leaders against the group's rating of them it was found that the Mean *D* score for the leaders was significantly lower ($p < .01$) than that for the non-leaders, i.e., were closer in their self ratings to the group ratings of themselves (Table 5). This was true for 20 of the 26 items. This supports the hypothesis that leaders are generally more realistic in their self assessment than non-leaders and, perhaps, are more insightful and more aware of group feeling.

TABLE 5
MEAN DIFFERENCE SCORES BETWEEN LEADER AND NON-LEADER SELF RATINGS AND GROUP RATINGS OF THESE SAME LEADERS AND NON-LEADERS FOR ALL ITEMS

	Leaders	Non-leaders
$\overline{D}$ Score	11.98	15.81
SD	3.67	3.66

D. SUMMARY AND CONCLUSIONS

Ten leaders and 10 non-leaders of social groups on the Queens College Campus were rated by nine members of their respective groups on a 26-item scale covering three major categories: Self Acceptance, Need Achievement, and Interpersonal Skills. The leaders and non-leaders also rated themselves on the same scale. A comparison was made between leader and non-leader mean ratings on each of the 26 items and a mean difference ($\overline{D}$) Score was obtained from the difference between group mean ratings and self mean ratings for leaders and non-leaders on the full scale.

In general, these results are co-ordinate with, and confirmatory of, the results of previous experiments (3, 4, 7, 11). The following conclusions seem warranted by the data:

1. In the sample studied the scale generally discriminated the leaders from non-leaders in that 23 of the 26 items gave differences in the expected direction.

2. Leaders of social groups are rated significantly higher than non-leaders in confidence and in degree of realism by the members of their groups.

3. Leaders of social groups are rated as being more willing to accept responsibility, and are regarded as being somewhat more driving, or arbitrary than non-leaders.

4. Leaders are rated as being more forceful and more persuasive than non-leaders. The interpersonal skill of dominance is balanced by the possession of tact, which is reflected in the high group ratings of the leaders on degree of diplomacy.

5. Leaders demonstrate, in their self-ratings, a significantly greater degree of awareness of how the group feels about them than do the non-leaders.

6. Three items (1, 13, 16) showed reversals from the previously hypothesized direction. Two (1 and 13) were statistically significant. That is, non-leaders were seen as "taking things in stride" more than leaders and as having greater "goal clarity" than leaders.

These reversals may reflect, in part, the fact that the leader of a group, to some extent, is an explorer or venturer into new fields. As such he might be perceived as having less clear goals and being more easily upset than non-leaders.

REFERENCES

1. BALDWIN, A. L., KALHORN, J., & BREESE, F. H. The appraisal of parent behavior. *Psychol. Monog.*, 1949, **63**, 26.
2. CHAMPNEY, H., & MARSHALL, H. Optimal refinement of the rating scale. *J. Appl. Psychol.*, 1939, **23**, 323-331.
3. CARTER, L. F., & NIXON, M. Investigation of the relationship between four criteria of leadership ability for three different tasks. *J. of Psychol.*, 1949, **27**, 245-261.
4. GIBB, C. A. The principles and traits of leadership. *J. Abn. & Soc. Psychol.*, 1947, **42**, 283.
5. GUETZKOW, H. (*Ed.*). Groups, leadership, and men. Pittsburgh: Carnegie Press, 1951.
6. HEMPHILL, J. K. Relations between the size of the group and the behavior of superior leaders. *J. Soc. Psychol.*, 1950, **32**, 11-32.
7. JENKINS, W. O. A review of leadership studies with particular reference to military problems. *Psychol. Bull.*, 1947, **44**, 75.

8. Leary, T. Interpersonal diagnosis of personality: A functional theory and methodology for personality. New York: Ronald Press, 1957.
9. Lewin, K., Lippitt, R., & White, R. K. Patterns of aggressive behavior in experimentally created "social climates." *J. Soc. Psychol.*, 1939, **10**, 271-299.
10. McClelland, D. C. The achievement motive. New York: Appleton-Century-Crofts, 1953.
11. Stogdill, R. M. Personal factors associated with leadership: A survey of the literature. *J. of Psychol.*, 1948, **25**, 35-71.

Department of Psychology
Queens College
Flushing 67, New York

Published as a separate and in *The Journal of Psychology*, 1959, **47**, 57-66.

LEADER ACCEPTANCE OF RESPONSIBILITY FOR GROUP ACTION UNDER CONDITIONS OF UNCERTAINTY AND RISK*

Fels Group Dynamics Center, University of Delaware

ROBERT C. ZILLER[1]

A. PURPOSE

Frequently, leaders are confronted with decision-making situations involving uncertainty and risk for the entire group, yet the circumstances do not permit the leaders to confer with their group members. For example, aircraft commanders occasionally must decide whether to attempt a crash-landing or order a bail-out. Unwillingness to accept responsibility for group action could result in a fatal delay or give undue weight to the course of action which relieves the leader of responsibility to a greater extent(bail-out in the example above). It is the purpose of this study to explore the relationship of several leader and group variables to leader action under a model situation of this kind.

B. SOME HYPOTHESES OF LEADER ACTION UNDER UNCERTAINTY

To the extent that "physical reality" is lacking, the basis upon which a decision is evaluated is "social reality," the degree to which the decision is supported by a reference group (4). Experimental evidence indicates that dependence upon "social reality" (social conformity), is, in turn, related to personality variables (8). Thus, it may be hypothesized, rather generally, that in an ambiguous decision-making situation, in which the group is involved but cannot be consulted, a leader's perception of the field and courses of action will necessarily be influenced by both personality and social variables.

The social variable selected for study was group attraction. In predicting the relationship between group attraction and acceptance of responsibility for group action a stochastic model of decision making (10) was invoked. The parameters involved are the *prize* which is the object which an individual stands to gain in the event of success; the *price* or stake which is the object which the individual stands to lose in the event of failure; the

---

*Received in the Editorial Office on October 6, 1958, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The data for this report were collected at Stead Air Force Base, Reno, Nevada, by personnel of the Survival Research Field Unit directed by Dr. E. Paul Torrance. The writer is indebted to Drs. Leonard Berkowitz and Thornton B. Roby who reviewed the original draft from which the present report developed.

probability of success which is the ratio of wins to the total number of trials. These parameters are regarded as psychological variables and are not necessarily related to real properties. Thus, the prize may be perceived as a gain in wealth or prestige. In the decision-making process, the price is compared with the prize under the condition of probability of success. Thus, under constant conditions of probability of success, the decision to act (gamble) may be predicted from the relative values assigned to the price and prize. Necessarily, under these conditions, if the price is decreased or the prize is increased, the favorability of the decision to act is increased.

With regard to group attraction, there is some evidence (7, 13) that leaders of high attraction groups are reinforced by their groups (their reference groups in a decision-making situation under stress) to a greater extent than leaders of low attraction groups. Assuming that increased social support reduces the perceived value of the stake or price or fear of failure under conditions of uncertainty, and that the prize and probability of success are constant, the ratio of the price to the prize is decreased and the perceived favorability of the decision to gamble (to assume responsibility for group action) is increased.

Three personality measures of the leader were selected for analysis: social conformity, the *R*-Scale, and motivation. In a laboratory experiment involving clinical ratings and task-criteria, Crutchfield (3) found that high conformists were self reliant and independent in judgment. The present experiment was designed to test this hypothesis in an experimental field setting.

However, a felt anxiety concerning the group's opinion may be incompatible with the leader's rôle concept. Thus, there is evidence (13) that leaders with high *F*-Scale scores (1) tend to arrogate responsibility for group action. Moreover, persons with high *F*-Scale scores in comparison with persons with low *F*-Scale scores tend to find ambiguity intolerable (1) and, thus, may be expected to initiate action more frequently in an ambiguous group decision-making situation.

Finally, it is hypothesized that persons highly motivated toward achievement (maintenance of a high standard of leadership in the present instance) are more likely to accept responsibility for group action. Again invoking a stochastic model of decision-making, this hypothesis evolves from the assumption that the relative perceived values of the prize (or goal) is greater for the highly motivated leader; and since, under these conditions, the ratio of the price to the prize is decreased, the perceived favorability of the decision to act is increased under constant conditions of probability of success.[2]

---

[2] A condensation of this report was read before the American Psychological Association convention, September 7th, 1955.

C. PROBLEM

Briefly the four hypotheses which this study was designed to test are: In an ambiguous group decision-making situation where circumstances do not permit the leader to confer with the group members, responsibility for group action is assumed more frequently by leaders (*a*) of relatively high attraction groups, (*b*) who are relatively unconcerned about the opinions of the group members, (*c*) with relatively high *F*-Scale scores, (*d*) with a high standard for personal achievement in the leadership rôle.

D. SUBJECTS AND EXPERIMENTAL PROCEDURE

1. *Subjects*

Thirty-nine B-26 aircrews comprising about 130 men took part in the experiment. For the most part, each crew was composed of three men of whom two were officers (pilot and navigator). However, some crews carried an additional airman. Leadership was the pilot's responsibility. All subjects had worked together as members of the same crew for approximately two months.

2. *Experimental Setting*

These crews were approaching the end of a training program designed to prepare them for combat in the Korean theatre of operations. In the phase of training during which this experiment was conducted, the crews were to learn how to survive in the event of a bail-out or crash landing anywhere in the world. During the initial days of this training, three or four crews were housed in the same tent. The tent also served as the testing room.

3. *The Decision-Making Situation*

Following a brief introduction to the experiment, the crew members completed a series of questionnaires which provided the data from which measures of the independent variables were derived. When the leaders turned to the last questionnaire[3] they were called outside the tent individually. Here, in an informal manner, the decision-making situation was described verbally to the leaders for their immediate consideration and decision. Every effort was made to make this individual decision-making situation appear as a task apart from the group decision-making session to follow. In this way, it was hoped that the questionnaire responses and decision would not be consciously related.

---

[3] An instrument intended to measure self esteem which was not included in this report after subsequent analysis (3) failed to validate what originally appeared to be a very satisfactory measure of this variable.

The decision-making situation involved a military intelligence exercise which was a renowned phase of the training program in which the men were participating, some of the details of which cannot be revealed for reasons of military security. In general, however, the exercise included an attempt by the crew members to evade "capture" in "enemy" territory for a given period of time. If captured, there was an unknown probability that individuals would be interrogated under conditions of extreme physical and psychological discomfort. The trainees had no way of estimating their chances of capture or their chances of being subjected to rigorous questioning under stress conditions.

Under the guise of assisting the military intelligence personnel in this training exercise, the leaders were offered a choice of action: (*a*) they could pursue the normal course of events in which case it was impossible to estimate how many of the crew would escape interrogation or (*b*) throw a die under the conditions that in the event a number 1, 2, 3, or 4, turned up, all crew members captured would be subjected to an intensely distressing experience, but in the event of a 5 or 6, all crew members who were captured would escape the interrogation.[4]

In their efforts to arrive at judicious decisions, the leaders frequently asked if they might consult their crew before submitting the final decision.[5] They also frequently rationalized their decision by stating that they would have made a different decision if the group members were not involved. Thus, there were indications that the situation assumed at least a degree of reality.

Another aspect of the problem should be clarified. It will be recalled that three or four groups were housed in the same tent. Assuming that the leaders foresaw an exchange of information concerning their respective decisions, a competitive element may have been introduced. However, it may be argued that this situational factor contributes to the validity of the model; competition between leaders and groups is commonly an element in group decision-making situations.

E. MEASURING THE VARIABLES

1. *Social Conformity*

Social conformity was measured by three items relevant to the respondents' concern about disagreeing with the group on an issue of vital importance to the group as a unit.[6] These items were presented in the following context:

---

[4] These conditions represent a one-third probability of winning. This probability was used in the experiment when pilot studies showed that one-half and two-thirds probabilities were accepted by all leaders of a sample population.

[5] Unfortunately, statements of this kind were not recorded as behavioral evidence of the leaders' concern with the opinions of the other members.

> Suppose you found yourself holding an opinion which was not in agreement with an opinion held by the other people in your crew and this opinion was on an issue of importance to both you and to the crew. In each set of statements below, check the one alternative which would best express how you would feel and what you would do if such a situation existed in your crew.
>
> 1. I would feel: (check one) (*a*) not at all concerned about my differing from the crew's opinion; (*b*) slightly concerned about my differing from the crew's opinion; (*c*) moderately concerned about my differing from the crew's opinion; (*d*) greatly concerned about my differing from the crew's opinion.
>
> 2. If I found that my opinion differed from that of this crew, I would feel: (check one) (*a*) just as sure of my own opinion; (*b*) less sure of my own opinion; (*c*) much less sure of my own opinion; (*d*) very much less sure of my own opinion.
>
> 3. If I found that my opinion differed from that of this crew, I would have: (check one) (*a*) a very strong desire to change my opinion to agree with that of the crew; (*b*) a strong desire to change my opinion to agree with that of the crew; (*c*) a moderate desire to change my opinion to agree with that of the crew; (*d*) hardly any desire to change my opinion to agree with that of the crew; (*e*) a desire to maintain my own opinion.

In calculating the total score, the alternatives of each question were assigned weights of 1 to 4 or 1 to 5, in accordance with the number of options. In this population the mean correlation of the items with the total score was .79.

2. *Group Attraction*

The index of group attraction follows directly from Festinger's definition of group cohesiveness as the resultant of all forces acting on the members to remain in the group (5, p. 164). The term "group attraction" is used here rather than the more general term "cohesiveness" as defined by Festinger, since the two questions involved in this index do not measure the forces acting to prevent the individual from leaving the group. These were ignored because of the involuntary nature of military groups.

The index was derived from the following questions (the same alternatives were used for each of the items):

> 1. If one man from each crew were ordered to remain at the home base rather than to attend Survival School and you were the man chosen,

---

[6] This instrument was adapted from a conformity scale originally developed by Alvin F. Zander, Research Center for Group Dynamics, University of Michigan.

how would you feel about this? (check one): (*a*) I'd feel very lucky; (*b*) I wouldn't mind; (*c*) It wouldn't matter one way or another; (*d*) I wouldn't like it, but would not feel too strongly about it; (*e*) I wouldn't like it a bit.

2. How would you feel about being transferred to another crew or about your crew being split up as replacements for other crews?

In calculating the index of group attraction, the alternatives of each item were weighted from 1 to 5, respectively. The score of each individual group member was the sum of the weighted responses. The mean score of the group provided the desired index. In an earlier study (13) an index of the reliability of the measure was calculated using a variation of the split-half technique. A Pearsonian correlation of .46 was obtained.

3. *F-Scale*

Only 28 of the original 30 *F*-Scale (1) items were used when it was observed that Air Force personnel were openly derisive regarding the one item, while another appeared outdated.[7] The usual six-point scale of agreement-disagreement was used. However, "Disagree very much" was assigned a weight of —3 while "agree very much" received a weight of 3. Reported reliability measures of the full instrument range from .81 to .91 with an average of .87 (1, p. 251).

4. *Motivation*

The measures of motivation, or need for achievement, were adapted[8] from an Ability Rating Scale (11) which was originally developed for this purpose. Subjects were required to rate themselves on the following abilities: (*a*) physical ability; (*b*) mechanical ability; (*c*) leadership ability; (*d*) social ability; (*e*) economic ability; (*f*) attracting and entertaining ability; (*g*) observational ability; (*h*) general intellectual ability; (*i*) scientific understanding; (*j*) psychological understanding; (*k*) artistic creative ability. Each ability was accompanied by a description. For example, mechanical ability was described as "the ability to manipulate mechanical appliances and instruments, to take apart, repair, and construct apparatus—electrical and mechanical. Technical skill in applied sciences."

---

[7] The items eliminated were: (*a*) "The wild sex life of the old Greeks and Romans was tame compared to some of the goings-on in this country, even in places where people might least expect it." (*b*) "It is best to use some prewar authorities in Germany to keep order and prevent chaos."

[8] Dr. Hugh B. Kohn of the Survival Research Field Unit, Crew Research Laboratory, Reno, Nevada, designed the drive measures used in the present study.

Two nine-point scales preceded each ability title. The men were required to indicate on the one scale how much of the ability they have and on the other how much ability they would like to have. The first scale (Motivation *A*) presumably involves a person's self-concept. The original authors (11) state that the scale probably indicated a standard which the individual felt compelled to maintain. The instrument, thus, was thought to measure a drive state of a rather complex nature.[9]

The second index (Motivation *B*) was derived from a summation of the weighted responses to the second scale and concerned goal aspiration. A third index (Motivation *D*) was obtained from the difference between the regressed scores from the actual ability index and the level of aspiration score.

The intercorrelation matrix of the independent variables is given in Table 1. On the basis of the high correlation (.92) between the independent variables Motivation *B* and *D*, the latter was not included in the final analysis.

TABLE 1

INTERCORRELATIONS OF THE INDEPENDENT VARIABLES

(N = 35 crews)

	1	2	3	4	5	6
1. Group attraction		—.17	—.22	—.08	—.18	—.20
2. Motivation (A)			.47*	.03	—.19	—.01
3. Motivation (B)				.92*	.49*	.06
4. Motivation (D)					.59*	.06
5. *F*-Scale						.14
6. Social conformity						

*Statistically significant at the .05 level of confidence or better.

F. RESULTS

In the sample of 39 aircrews, only nine leaders elected to throw the die. Thus, while the results are all in the expected direction, only the results with regard to social conformity approach the accepted standard for statistical significance (see Table 2). Nevertheless, the results are suggestive and warrant brief comment.

The results suggest that a leader who is relatively unconcerned about the opinions of the group members, relatively "authoritarian," and the leader who has a high level of aspiration tends more readily to assume responsibility

---

[9] This assumption was validated by correlating scores on the scale with other motivational variables identified by Murray, *et al.* (9).

a. Need achievement	—.65
b. Need dominance	—.79
c. Need exhibition	—.69
d. Ego ideal	—.65

TABLE 2
DIFFERENCE BETWEEN GROUPS WHOSE LEADERS ASSUME RESPONSIBILITY FOR GROUP ACTION

	Leader action			Leader inaction				Level of significance*
	Mean	*N*	*SD*	Mean	*N*	*SD*	*t*	
1. Group Attraction	9.09	9	0.637	8.90	30	0.804	.64	.26
2. Motivation (A)	61.22	9	8.702	58.48	29	8.207	.84	.20
3. Motivation (B)	88.22	9	8.107	84.10	29	9.557	1.15	.13
4. *F*-Scale	9.50	8	20.131	— 1.81	27	22.101	1.28	.10
5. Social Conformity	6.56	9	1.571	7.53	30	1.615	1.54	.06

*One-tailed probability.

for group action. These characteristics of the action-seeking leader are similar to those described by Crutchfield (in a report cited at the outset) as being related to nonconformity (3). In contrast with the high conformists, the extremely independent men were described more often by an independent assessment staff as effective leaders, ascendent in their relations with others, and self reliant and independent in judgment. Thus, in some respects and to a limited extent the findings presented here may be viewed as confirmation in a field situation of Crutchfield's laboratory results.

However, the constellation of relationships takes on added meaning if the decision-making situation is conceptualized as a risk-taking situation. Viewed in this way, the "high risker" (with regard to the prize of avoidance of punishment for the group members) is described by the independent variables as independent, highly motivated, and authoritarian. The results then appear to corroborate a recent study of biographical correlates of risk-taking tendencies (12) in which it was revealed that the high-risker was self-confident, physically and socially adequate, competitive and self-expressive.

At another level of abstraction, however, the findings of the present study along with the two studies cited above, suggest that the element common to leaders who accept responsibility for group action, non-conformists, and persons with high risk-taking tendencies is the existence of a reservoir of personal and social resources sufficient to provide a wide tolerance in the event of misfortune. Or in terms of decision theory, the value of the stake is relative to the resources available. Consequently, under constant conditions of probability of success and stake or prize, the decision to act will be made more frequently by the persons with greater resources at their disposal.

To the extent that action is preferred to inaction by the group members under conditions of uncertainty and stress, such leader characteristics which tend to relate to willingness to accept responsibility for group action appear to be positively weighted. However, evidence has been accumulated which indicates that leadership is related to the situation (6). Thus, while in the present situation, high *F*-Scale leaders tended to assume responsibility for group action under conditions of uncertainty and risk; in a problem-solving situation (13) military groups in which the leaders had high *F*-Scale scores were found to be less adaptable.

G. SUMMARY

This experiment was designed to explore the characteristics of leaders who are willing to assume responsibility for group action under conditions of uncertainty with regard to group security. The independent variables included measures of group attraction and the leader's social conformity, *F*-Scale score, and motivation. The decision-making incident was set in a realistic military training situation in which rather severe physical and psychological discomfort might result. In essence, the leader was required to choose: (*a*) to assume responsibility for group action by throwing a die to determine whether or not the group would avoid a threatening group experience, or (*b*) to proceed in the customary manner (customary in that previous training groups and their leaders had been given no choice) against unknown odds under circumstances created by chance events and individual behavior.

In this choice situation 9 of 39 leaders accepted the responsibility for group action. These men tended to be relatively unconcerned about differing from the opinions of the group, scored relatively high on the *F*-Scale, and were generally more highly motivated.

REFERENCES

1. ADORNO, T. W., *et al.* The Authoritarian Personality. New York: Harper, 1950.
2. COHEN, A. R. A report on some explorations of self-esteem: Progress report (mimeographed), August, 1954. Research Center for Group Dynamics, University of Michigan, Ann Arbor, Michigan.
3. CRUTCHFIELD, R. S. Conformity and character. *Amer. Psychol.*, 1955, **10**, 191-198.
4. FESTINGER, L. Informal social communication. *In* L. Festinger, *et al.*, Theory and Experiment in Social Communication. Ann Arbor, Michigan: Edwards, 1950.
5. FESTINGER, L., SCHACHTER, S., & BLACK, K. *Social Pressures in Informal Groups: A Study of a Housing Project.* New York: Harper, 1950.
6. GIBB, C. A. Leadership. *In* G. Lindzey (*Ed.*), *Handbook of Social Psychology,* Cambridge: Addison-Wesley, 1954.

7. HARTLEY, E. L., & HARTLEY, R. E. Fundamentals of Social Psychology. New York: Knopf, 1952.
8. JANIS, I. L. Personality correlates of susceptibility to persuasion. *J. of Person.*, 1954, **22**, 504-518.
9. MURRAY, H. A., *et al.* Explorations in Personality. New York: Oxford Univ. Press, 1938.
10. PRESTON, M. G., & BARRATTA, O. An experimental study of the auction-value of an uncertain outcome. *Amer. J. Psychol.*, 1955, **61**, 183-193.
11. SMITH, R. G., JR. Self-rating of abilities as a measure of drive. Research Note 52-9, 1952, Air Force Personnel and Training Research Center, Lackland Air Force Base, San Antonio, Texas.
12. TORRANCE, E. P., & ZILLER, R. C. Risk and life experience. Lackland Air Force Base, Texas: Air Force Personnel and Training Research Center, February, 1957. (Research Report, AFPTRC-TN-57-23, ASTIA, Document No. 07892 C.)
13. ZILLER, R. C. Leader flexibility and group cohesiveness: determinants of group problem-solving processes and concomitant affective group member behavior. *Amer. Psychol.*, 1953, **8**, 459 (Abstract).

Fels Group Dynamics Center
University of Delaware
Newark, Delaware

Published as a separate and in *The Journal of Psychology*, 1959, **47**, 67-80.

A STUDY OF THE BRAID EFFECT: (HYPNOSIS BY VISUAL FIXATION)*

Department of Psychology, Stanford University

André M. Weitzenhoffer, Philip B. Gough, and Judah Landes[1]

A. Introduction

In 1841 James Braid (2) described a condition of "nervous sleep," or "neuro-hypnotism," as he called it, which could be brought about in five to ten minutes through the simple expedient of having a person fixate his eyes (and attention) upon a small bright object in such a way as to cause "a double upward internal squint." Because many of the phenomena he described have been believed to be closely related to those found in the more familiar type of hypnosis as we know it today, it has seemed desirable both to see whether his phenomena could be replicated and what their relationship is to the phenomena of hypnosis as induced with the aid of verbal suggestions.

The syndrome of neuro-hypnosis or, as we shall call it, the Braid effect, appears to be defined by eight symptoms (2, page 109f) as follows, these symptoms occurring in the order given:

1. *Pupil size:* First a constriction (due to consensual reflex), then a dilation.

2. *Pupil motion:* Braid speaks of a "wavy" motion of the pupil without further details.

3. *Eyelid closure:* May occur spontaneously. If not, the bringing forward and downward by the experimenter of the extended and slightly separated fore and index finger, from the fixated object to the *S*'s eyes will usually cause eye-closure. Eyelid closure tends to be spasmodic and the eyelids show a vibratory motion. In case this last procedure fails to cause the eyes to close, one then tells the *S* to allow his eyes to close the next time this is done, but to maintain his eye balls in the same position and his mind riveted on the thought of the object.

---

*Accepted for publication by Ernest R. Hilgard of the Editorial Board, received in the Editorial Office on October 12, 1958, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The authors wish to acknowledge the assistance of Dr. Ernest R. Hilgard, director of the program under which this investigation was conducted. This is one of a series of studies from the Laboratory of Human Development, supported by a grant from the Ford Foundation.

4. *General tonic rigidity:* Body members resist passive movement.

5. *Specific catalepsies:* Primarily in the form of waxy flexibility, observed when manipulating a member with rigidity setting in. The latter may spread.

6. *Pulse rate:* Rises markedly following the production of a specific catalepsy.

7. *Hypersuggestibility:* Can be demonstrated by describing to some other person fictitious effects which are to be elicited in the *S*. Thus, Braid demonstrated the various presumed (by the Mesmerists) effects of magnets with a piece of wood by describing these effects to an observer near the hypnotized *S*.

8. *Stupor:* If the Braid effect is allowed to persist without further stimulation it becomes transformed into a stuporous state in which the *S* becomes unresponsive.

Although phenomena exhibiting various "hypnotic" features were described much earlier, it is generally agreed that modern, scientific hypnotism begins with Braid's publication. There is, however, some question whether the syndrome Braid described, and thereafter referred to as the Braid Effect, is the same thing or is even closely related to that which is currently called "hypnosis." As a matter of fact, around 1886, Bernheim (1) even questioned the reality of the Braid Effect, taking the position that all hypnotic phenomena, including "hypnosis," as then understood, were suggested artifacts. "There is no hypnosis," said he, "but only suggestion" (1, p. 47). But Bernheim never gave a truly satisfactory proof of the correctness of his contention. True, he showed he could bring about the various symptoms and phenomena of hypnosis by suggesting them, but this cannot prove that such phenomena cannot exist independently of suggestion. Although suggestion has been given an increasingly important place in the study and production of hypnosis and hypnotic phenomena, visual fixation has, somewhat paradoxically, been retained as a basic element in many modern procedures for inducing hypnosis. Because of the prevalent use of eye fixation one ought to ascertain in a more definite manner whether or not the technique devised by Braid brings about any kind of effect independently of suggestion, and if so, what the relationship of such effects to modern hypnosis really is. It has been the main purpose of the present investigation to ascertain whether or not visual fixation alone, used in the manner Braid used it, can bring about the condition he described or in any event a condition which can be identified with our present day concept of hypnosis.

B. PROCEDURES

1. *Subjects and Groups*

In order that our *S*s might be ignorant of the nature of the experiment (both of the possible existence of the Braid effect and of its relation to hypnosis) they were invited to the laboratory to participate in an experiment on "attention." All *S*s were members of the class in introductory psychology.

In order to produce the Braid effect the *S* was asked to fixate a small bright button binocularly for 10 minutes. Presumably his eye-movements were being recorded as he attempted to sustain attention for this time. This arrangement followed very closely an experiment of Braid (2, p. 100f). At the end of the 10-minute period a series of tests were made (to be described in detail later). According to Braid, pronounced symptoms should be produced through such visual fixation, without any accompanying suggestions.

In order to determine the relationship of such effects as appeared (or failed to appear) to hypnosis as induced by standard procedures,[2] the *S*s completing the "attention" experiment were asked back to the laboratory for a second time. At that time they were given some waking suggestions (postural sway) and then invited to undergo a brief hypnotic induction. The susceptibility to hypnosis was tested by a modified form of the Friedlander-Sarbin scale (4).

In all, six groups of *S*s were run under somewhat different conditions. These groups are summarized below.

Group I. Minimal instructions for Braid effect (12 males; 14 females). The sole instructions and information these *S*s received were that they were to concentrate and keep their eyes and attention upon the target until told that the experiment was over.

Group II. Permissive about eye closure (14 males, 13 females). The *S*s received essentially the same information and instructions as those of Group I, but worded in a highly permissive manner aimed at allowing the *S* freedom to produce whatever responses he might feel inclined to produce.

Group III. Suggested expectation of Braid effect (16 males, 9 females.) The *S*s were told in a very authoritative manner that certain effects were expected to take place, and what these effects were. They were again told to allow the occurrence of whatever effects they felt were impending.

For each of these groups the following program was followed:

Day 1. Braid effect.

---

[2] Such as the one described by Friedlander and Sarbin (4) which was used in the present investigation.

Day 2. Waking suggestion followed by standard hypnotic induction[3] (about one week later).

Group IV. Minimal instructions; modified test for Braid effect; added suggestibility tests (16 males, 5 females).

Day 1. Braid effect tested prior to eye closure; suggestibility tests from Friedlander-Sarbin scale.

Day 2. Omitted.

Group V. Minimal instructions; added Friedlander-Sarbin Suggestibility tests (16 males, 5 females).

Day 1. Arm catalepsy and suggested arm rigidity tests for the Braid effect used only if involuntary eye closure takes place; suggestibility tests from Friedlander-Sarbin scale.

Day 2. Omitted.

Group VI. Volunteered for "experiment in hypnotism"; expected Braid effect (15 males, 13 females).

Day 1. Braid effect tested after eye closure; suggestibility tests from Friedlander-Sarbin scale.

Day 2. Waking suggestion followed by standard hypnotic induction.

Thus a total of 148 *S*s, 89 men and 59 women, served in the experiment. All were at the start quite unsophisticated about hypnotism, and only one had ever been hypnotized before.

2. *Apparatus and Specific Procedures*

a. Induction of the Braid Effect. Each *S* participating in the first part of the experiment was led to a small room and asked to seat himself comfortably in an upright upholstered chair. The *S* was then told that we were interested in studying eye-movements and in obtaining some physiological measures such as pulse rate during intense fixation of the eyes and attention upon a target. The target consisted of a small flat, disk-like, shiny, metallic button 0.12 inches in diameter placed against the white background of a posterboard 10″ x 14″ at a distance of 15 inches from the eyes. The lens of a moving pictures camera could be seen protruding through an opening in a large masonite partition facing the *S* who was told that this camera would be used to film his eye movements. This instrument, merely used as a prop, was "operated" by a second experimenter sitting behind the partition, of whom the *S* was aware. The position of the target was adjusted in such a way as to force the *S* to look upward at the target, with some strain involved. The eight symptoms summarized earlier were then sought, forcible closure of the eyes being brought about at the end of 10 minutes in accordance with Braid's instructions in the event that the *S* did not close his eyes spontaneously. At the termination of this portion of the experiment, each *S* was asked to go to an adjoining room for a short interview with the second experimenter (who had manipulated the camera). This interview, which usually lasted about 15 minutes, aimed at obtaining information with regard to the *S*'s subjective experiences and his perception of the experiment. At the end of the interview the *S*s were asked not to communicate with others about the experiment until it was over.

---

[3] Of the 78 *S*s participating in the Braid effect portion of the experiment, four failed to participate in the hypnotic portion, one each from Groups I and II, two from Group III.

As already indicated we introduced various elements of set, expectation, or suggestion in the course of studying the Braid effect. With the exception of the last group of *S*s this was done by giving the subjects different instructions and information just prior to starting visual fixation, Group I receiving no clues regarding what we really were studying, whereas Group III received a maximum of clues and suggestions with regard to the Braid effect.

b. Postural Sway Test and induction of hypnosis. All *S*s participating in the first part of the experiment (that is, Groups I through III) were asked to report within a week for a second session ostensibly having to do with a further study of attention. This took place in a different room. The session began with each *S* being told that now we wanted to study the possible effects that concentrating on body movements might have upon body posture. The *S* was then asked to stand eyes closed, head level, heels and toes together and hands by his sides, and to pay close attention to what was to be said. The experimenter standing close by and behind the *S* then gave approximately one minute[4] of suggestions of backward sway and fall. Appreciable swaying without falling was scored 1, falling received a score of 2. Following this, the *S* was asked to sit down in an upright upholstered chair and it was explained to him that since it was said that concentration of attention was involved in the production of hypnosis we thought this would be a nice occasion to study this possible aspect of hypnotism and wondered if the *S* would mind being hypnotized. If he acquiesced, a thumb tack stuck in the ceiling at about 8 feet distance from the *S*'s eyes was pointed out to him and he was asked to fixate his eyes upon it. The remainder of the induction followed the pattern and used essentially the same words given by Friedlander and Sarbin. The same set of test suggestions as given by these authors was also used with the addition of a suggested visual hallucination.[5] Scoring followed that used by Friedlander and Sarbin, the hallucination being scored 1 if there was evidence of a weak hallucination, and 2 if it was a strong hallucination.

---

[4] The same words were used by all experimenters, read from a typed sheet. In all four different operators were involved in this part.

C. EVIDENCE FOR THE BRAID EFFECT

We have defined the Braid effect as the presence of eight symptoms described by Braid, these symptoms occurring on the basis of visual fixation alone, unaccompanied by verbal suggestions or expectation of hypnotic effects. In so far as these symptoms are concerned, none of our *S*s exhibited the Braid effect *in toto*. The most striking finding in this regard is that but a single *S* closed his eyes spontaneously during the 10-minute fixation period, and he was the one who had been previously hypnotized. Thus, what Braid had described as a highly characteristic sign failed to appear. On the other hand, some relevant and scorable accompaniments of prolonged eye fixation were observed: spontaneous *arm catalepsy* (waxy flexibility), *indirectly suggested arm rigidity, sleepiness, spontaneous hallucinations, a feeling of dissociation*, and *lethargy*. Table 1 shows the frequency of occurrence of these six "symptoms" and of involuntary eye closure for Groups I, II, and III.

TABLE 1
BRAID EFFECT SYMPTOMS

Symptoms	Number of subjects with symptoms Group I ($n = 26$)	Group II ($n = 27$)	Group III ($n = 25$)	Total ($n = 78$)	Per cent
1. Sleepiness	14	11	12	37	47
2. Indirectly suggested arm rigidity	9	12	11	32	41
3. Spontaneous arm catalepsy	7	6	5	18	23
4. Spontaneous hallucinations	4	1	4	9	12
5. Lethargy	2	2	3	7	9
6. Feeling of dissociation	0	1	3	4	5
7. Involuntary eye closure	1	0	0	1	1

Note: Per cents add to more than 100 because some subjects gave more than one symptom.

As can be seen, arm catalepsy and arm rigidity, the two manifestations most like those Braid reported which we observed, occurred in 23 and 41 per cent of cases respectively. It may also be noted that the various modifications of instructions within the three groups seem to have had little effect so far as producing the observed phenomena. For the purposes of further analysis we defined a "Braid score" for each subject as the number of symptoms listed in Table 1 which he exhibited. This gave a score with a possible range of 0 to 7. It may be remarked in passing that none of the *S*s scored above 4. A Kruskal-Wallis one-way analysis of variance was performed on the Braid scores obtained by the three groups in order to test the null hypothesis that we were dealing with the same population in all three instances. Although there conceivably was some slight influence of the differing conditions ($.05 > p > .02$) we felt justified in pooling the results from the three groups for further analysis.

---

[5] This is the set we shall later refer to as the "hypnotic test suggestions."

Obviously, while we cannot say that we have observed the full Braid effect certain definite effects were found to be present in some subjects in conjunction with prolonged visual fixation. Whether visual fixation was specifically responsible for the appearance of these effects we cannot say. Conceivably the same individuals would have shown them had they simply been asked instead to sit with their eyes closed and their minds blank for 10 minutes. Unfortunately we did not try this.

D. RELATION BETWEEN BRAID EFFECT, RESPONSE TO VERBAL SUGGESTIONS, AND SUSCEPTIBILITY TO HYPNOSIS

On the assumption that visual fixation had something to do with the incidence of the symptoms of Table 1, the next question we want to examine is to what extent *S*s who were and were not affected by visual fixation alone also were or were not affected by (*a*) suggestions of postural sway given to them in the absence of a prior hypnotic state, and (*b*) modern procedure aimed at inducing hypnosis. The distributions of scores with respect to Braid effect, postural sway, and hypnotic susceptibility for the pooled results of three groups are presented in Table 2.

TABLE 2
SCORE DISTRIBUTIONS FOR BRAID EFFECT, POSTURAL SWAY, AND HYPNOTIC SUSCEPTIBILITY (GROUPS I-III)

Braid effect (possible score = 7)		Postural sway		Hypnotic susceptibility (possible score = 22)	
Score	Cases	Score	Cases	Score	Cases
0	23	0	25	0- 5	24
1	23	1	12	6-10	15
2	16	2	41	11-15	19
3	12			16-20	9
4	4			20+	7
Total	78		78		74

Note: The lesser number of *S*s receiving susceptibility score was caused by the fact that 4 *S*s declined to be hypnotized.

When the data are rearranged in the form of the two contingency tables of Tables 3 and 4 it is found that there is a small but statistically significant positive relationship between Braid scores and hypnotic susceptibility, whereas there is no statistically significant relationship between Braid scores and postural sway scores. This would seem to suggest that something other than waking suggestibility as measured by postural sway is involved in the production of the phenomena we observed in conjunction with visual fixation and that this element is common to the situation defined by the modern procedure we used to induce hypnosis.

TABLE 3
ASSOCIATION BETWEEN BRAID EFFECT AND HYPNOTIC SUSCEPTIBILITY (GROUPS I-III)

	Susceptibility	
Braid score	0-15	16-22
<3	50	9
$\geqslant 3$	8	7

$N = 74$, $df = 1$.
$x^2 = 6.96$, $C = .29$.
$.01 > p > .001$.

TABLE 4
ASSOCIATION BETWEEN BRAID EFFECT AND POSTURAL SWAY (GROUPS I-III)

	Postural sway	
Braid score	0	1-2
<3	21	37
$\geqslant 3$	2	14

$N = 74$, $df = 1$.
$x^2 = 3.29$, $C = .21$.
$.10 > p > .05$.

It had been hoped that a study of the *S*s' responses to the test suggestions used in conjunction with the determination of hypnotic depth following 10 minutes of visual fixation would throw additional light on the relationship between visual fixation and hypnosis. To this end, Group IV was introduced into the study. Spontaneous arm catalepsy turns out to be the only symptom we can use to establish a Braid score for the subjects of Group IV. Thirty-three per cent of *S*s showed this symptom for this group, as against 21 per cent for the combined first three groups. This difference of 12 per cent is not statistically significant at the 5 per cent level of confidence and we feel we can reasonably consider Group IV to have been comparable to the first three groups. No *S* scored higher than 4 on the hypnotic test suggestions,[6] and since no *S* scoring less than 3 exhibited any signs of amnesia, it seems reasonable to use a suggestibility score of 3 as cutting score for the purpose of setting up the 2 x 2 table of Table 5. As can be seen the Chi square value obtained from it is not significant and we must conclude that no relation has been demonstrated between the production of spontaneous arm catalepsy and the exhibition of suggestibility of the type studied in connection with modern hypnosis.

---

[6] It should be remembered in this connection, that for obvious reasons, these scores

TABLE 5
SPONTANEOUS ARM CATALEPSY

Suggestibility (Friedlander-Sarbin Scale)	Present	Absent
$\leqslant 2$	3	9
$\geqslant 3$	5	4

$N = 21$, $df = 1$.
$x^2 = 2.04$.
$.20 > p > .10$.

D. EXPECTATION OF HYPNOTISM AND THE BRAID EFFECT

When the *S* expects to be hypnotized, the procedure of prolonged visual fixation heightens his suggestibility. This is evident from the results of Group VI, in which the *S*s volunteered for a hypnotic experiment, compared with those of Groups IV and V in which there was no expectation of hypnosis. A Mann-Whitney *U*-Test performed on the raw data shows that the difference in performance of Group IV and of Group V was not statistically significant, which fact enables us to pool these two groups. The comparison (Table 6) between the performance of the combined groups and that of Group VI differs in a statistically significant manner at better than the 1 per cent level of confidence (Mann-Whitney U-Test on the raw

TABLE 6
HYPNOTIC SUGGESTIBILITY FOLLOWING VISUAL FIXATION WITH AND WITHOUT EXPECTATION OF BEING HYPNOTIZED

	Number of *S*s			
	No expectation Groups IV and V		Expectation Group VI	
Suggestibility	Cases	Per cent	Cases	Per cent
0- 2	37	88.1	11	41
3- 5	3	7.1	7	26
6- 8	1	2.4	7	26
9-11	1	2.4	2	7
Total	42	100	27	100

---

do not include points for the time taken by the eyes to close as do the scores used in measuring depth of hypnosis when the latter is induced by the Friedlander-Sarbin method. Had this not been the case the highest score could have been 9.

data). Thus we may conclude that suggestibility, as ascertained on a standard scale normally used to measure hypnotic suggestibility, is higher when visual fixation is accompanied by the information that it will induce "hypnosis" as contrasted to the situation where such information is not available.

This, of course, says nothing as to how visual fixation affects suggestibility as compared to the effects on suggestibility of a standard induction of hypnosis. For the answer to this question we must look at another set of data obtained from Group VI. As may be seen from Table 7 when the *S*s of this group returned to the laboratory to be hypnotized by the Friedlander-Sarbin method, they scored essentially the same as they did following prolonged fixation. A Wilcoxon matched pairs test failed to reveal any difference between the two sets of scores ($p > .05$).

TABLE 7
HYPNOTIC SUGGESTIBILITY FOLLOWING VISUAL FIXATION ALONE AND WITH TRANCE INDUCING SUGGESTIONS (GROUP VI)

	Fixation plus Expectation		Standard hypnotic procedure	
Suggestibility	Cases	Per cent	Cases	Per cent
0- 2	11	41	14	52
3- 5	7	26	6	22
6- 8	7	26	3	11
9-11	2	7	4	15
Total	27	100	27	100

The Spearman rank correlation is rho = .72 and significant at the .01 point. This is an appreciable correlation, which suggests that when the *S* expects hypnosis visual fixation alone can lead to a hypnotic state; nevertheless we cannot ignore the question of what residual differences there are between induction of hypnosis through visual fixation alone and through visual fixation accompanied by verbal suggestions of relaxation and sleep. Unfortunately we do not have comparable data with regard to the induction of hypnosis without visual fixation when a *S* merely sits in a chair expecting to become hypnotized.

E. DISCUSSION

Reading the works of Braid (2), Durand de Gros (3), and other practitioners of "Braidism," the production of the Braid effect appears to be quite easy, and the effects to be clear-cut. Our experience shows that while this may have been true once upon a time, it was not so in our case. Neither do our results entirely support the contentions of the Nancy School (headed by Bernheim) that the Braid effect was solely the result of suggestions. We may assume that the effect existed and if so it may be of interest to find under what conditions it can be elicited. Obviously this is another problem. Ours was to produce the effect under conditions which closely approximated those under which the "Braidists" operated. We first assumed that suggestions, expectation, etc., were not important and obtained no effect. We then turned around and introduced various suggestions, mild to strong, and still there was no effect. We feel that this much we have shown: (*a*) if the Braid effect exists, it is not a universal phenomenon in our times. Perhaps it is a function of age, class status, education, or what not, but it certainly cannot be brought about easily in a typical Stanford University undergraduate population; (*b*) visual fixation *per se* seems to produce certain effects in many individuals, but in the absence of a rather specific structuring of the situation in which it takes place hypnosis does not ensue in its usual form.

The question of the nature of the *S* sample as a factor is perhaps worth discussing a little further. One of the most striking characteristics of our 148 *S*s is that with one exception they persisted in keeping their eyes open and upon the target until told to close them. There is nothing remarkable in this persistence even though this was uncomfortable and many *S*s would have preferred to close their eyes. What is striking is that our original instructions seem to have been interpreted by the *S*s (as stated in the interviews) as requiring them to keep their eyes on the target no-matter-what-happened.

Actually, even if this were the *S*'s understanding, it should not have prevented the Braid effect from taking place as there is little question from Braid's accounts and those of his students that their subjects were usually emphatically told to maintain their eyes at all cost upon the fixation object. Nevertheless, it seemed to us that we should do everything possible to promote the effect and consequently Group II was introduced. Short of suggesting the specific symptoms, the *S*s were told to do what they felt most like doing at any time, including closing the eyes and taking a nap. But even in this situation, and for reasons which escape us, those among the *S*s who felt like closing their eyes but did not, still seem to have felt that the instruction to keep their eyes upon the target was implicit, and they allowed this to dominate everything else. Was it the fact that we were presumably taking pictures of their eye movements and did they fear that closing their eyes would interfere with this? This possibility does suggest itself, although no subject spontaneously mentioned it. When this possibility was brought up in some instances, the answers have been noncommittal. In fact when directly asked how one might have presented the experiment to them so as not to give the impression that the eyes must be kept open at all cost none of the *S*s had any ideas to offer.[7] They were as vague about this as they were about why they felt they had to keep their eyes open even when they were inclined to close them. If students are so motivated to carry out instructions to keep their eyes upon the target in spite of strong suggestions to the contrary why is it that they do close their eyes when the Friedlander-Sarbin technique is used? The initial instructions then are to keep the eyes on the target "as long as you can." A few, but very few subjects in the standard hypnotic situation report not closing their eyes because of this instruction. There is a factor here which bears further study. Just what it is remains unclear, but a reasonable guess is that it is related to the specific interpersonal relationship which is established between subject and hypnotist at the time hypnosis is induced or even the possibility of doing so is brought up.

Another feature of our results which might be called to attention are the spontaneous hallucinations which some *S*s experienced. One *S*, for instance, saw many eyes looking at her. Another saw her brother's face. In many instances the hallucinations seem to have started out more as an illusion built around a distorted perception of the target. This sort of thing brings to mind the recent reports on sensory deprivation. It has indeed been suggested that hypnosis is the same as the condition brought about by decreased stimulation. It seems plausible that the Braid effect or something akin would be more directly related to states induced by sensory deprivation than to modern hypnosis in which after all, in spite of intense concentration upon bright objects and the hypnotist or his words, there is rather considerable and varied stimulation. In any event, it should be kept in mind that spontaneous hallucinations and other spontaneous manifestations are not infrequently seen to take place during the induction of hypnosis by modern methods and consequently our observations are not specific only to the Braid effect or the situation going along with its production.

Our results do not lead one to expect much contribution to the hypnotic state from visual fixation *per se,* but there remains to be seen how important a part it plays when combined with trance inducing suggestions. It is possible that it serves a very useful, although not essential function, as a tool for focusing and narrowing attention prior to it being channelized toward the hypnotist, assuming this last is an essential factor in the induction of hypnosis. A certain amount of important research remains to be done in the area of attention in order to understand fully what hypnosis is.

---

[7] In spite of the fact it might well have been an atypical case, one should point out here that in one experiment Braid hypnotized a servant of his in less than 10 minutes under conditions such that the motivation to keep his eyes open must have been very strong. This was, incidentally, a situation in which the subject did not know he was being "hypnotized."

Lastly it seems worth pointing to an intriguing if poorly understood feature of our results. Direct and indirect suggestions, weak and strong, of *specific* symptoms of the Braid effect had no influence. On the other hand, a very *general* and really quite vague suggestion implicitly contained in statements made to the *S*s telling them they were to be "hypnotized" had a definite influence, in spite of the fact that the subjects had initially no clear ideas as to what hypnosis was like. What is the factor which makes the difference? Is it the implied interrelationship? Is it the ambiguity inherent in the situation? Only further research can tell.

F. Summary and Conclusion

1. Volunteer undergraduate students (148 in all) were requested to fixate their attention and eyes for 10 minutes upon a small bright object in accordance with a procedure described by James Braid for the production of hypnosis. The *S*s were apportioned among six groups and accordingly received different information and instructions prior to and during the experiment. The instructions led 120 *S*s to believe that the experiment was a study of attention. They made up five groups for which the experimental conditions were varied in degree of permissiveness regarding what the subject could or could not do, with regard to information about what we expected to see happen, and the extent to which these effects were suggested. In addition, several groups were given a set of standard suggestions following forced closure of the eyes at the end of the fixation period. The sixth group of 28 *S*s believed the experiment was a study in hypnotism aimed at comparing methods of induction. Four of the groups also submitted knowingly at a later date to a standard induction of hypnosis by means of a modern method combining visual fixation and sleep-suggestions.

2. The full Braid effect as defined by eight symptoms reported by Braid was never obtained, even in the presence of strong suggestions. Evidence that visual fixation alone brings about heightened suggestibility related to hypnosis was, however, observed.

3. Specific suggestions of and information with regard to the symptoms of the Braid effect had no noticeable effects. In contrast, however, the relatively ambiguous information that the method will induce hypnosis leads to an appreciable over-all increase in suggestibility.

4. It is concluded that, in at least the type of population the Stanford University undergraduates who volunteered for our experiment represents, visual fixation *per se*, that is, without preparatory set or accompanying suggestions does not produce hypnosis in a form comparable to that brought about by modern procedure.

REFERENCES

1. BERNHEIM, H. Automatism et Suggestion. Paris: Alcan, 1917.
2. BRAID, J. Neurypnology: or the Rationale of Nervous Sleep Considered in Relation with Animal Magnetism. London: Churchill, 1843.
3. DURAND DE GROS. Cours Théoriques et Pratiques de Braidisme ou Hypnotisme Nerveux. Paris: Baillière, 1860.
4. FRIEDLANDER, J. W., & SARBIN, T. R. The depth of hypnosis. *J. Abn. & Soc. Psychol.*, 1938, **33**, 281-294.

Department of Psychology
Stanford University
Stanford, California

Published as a separate and in *The Journal of Psychology*, 1959, **47**, 81-86.

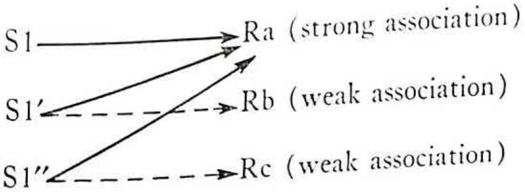
SOME EFFECTS OF MEPROBAMATE ON HUMAN LEARNING*

Department of Psychology, University of Michigan

EUGENE BURNSTEIN AND DONALD DORFMAN[1]

A. INTRODUCTION

Some recent studies (2, 6, 11, 12) have provided evidence for the hypothesis suggested by Hull's theory that in a complex memory task with large inter-item competition, a relatively high level of anxiety or emotionality will interfere with the speed of learning. With verbal material inter-item competition can be produced by taking advantage of certain properties of languages. Past experience produces a tendency for words to be associated with other words to varying degrees. In paired associate learning, this fact may be utilized to control the degree to which each stimulus word will elicit its own paired word and the degree to which it will tend to elicit response words that differ from the one with which it has been paired. If, as in the present experiment, the task consists of learning a list of 12 paired associate word items, high inter-item competition would be produced in the following manner: Four stimulus-response pairs of high associative connections are constructed. Two other stimulus words are selected which are synonymous with each of the four original stimulus words. These eight stimulus words are paired with response words with which they have minimal associative strength. Thus, in these four sets of three synonymous stimulus words, one is paired with a highly associated word and two are paired with weakly associated words. This may be diagramed as follows:



---

*Received in the Editorial Office on October 17, 1958, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The authors are indebted for valuable discussion and assistance to Dr. Leonard Uhr and Dr. James G. Miller, of the Mental Health Research Institute, Department of Psychiatry, and to the Behavioral Science Research Project of the University of Michigan, through which funds were secured.

Since S1′ and S1″ are synonymous with S1, they initially have a high associative connection with Ra. Thus in learning S1′ – – – Rb or S1″ – – – Rc, a relatively strong competing response tendency would be involved. More errors occur in learning such a list than is the case when stimulus words are dissimilar and competing response tendencies, thus, are weak.

In terms of Hull's theory a high degree of anxiety or emotionality directly reflects the drive level (D); a heightened drive level is said to increase both the number of competing responses which are above threshold and the differences in reaction potential between any dominant incorrect responses and correct responses, i.e., $E_a - E_b = D\ (H_a - H_b)$ (9, 10). It follows that highly anxious or emotional individuals should have a more difficult time in learning a list of paired associate words with high inter-item competition than individuals with a relatively low level of anxiety.

In the studies which supported this hypothesis, anxiety level has been manipulated by selecting individuals in the upper and lower 20 per cent of the scores on the Taylor Scale of Manifest Anxiety (12). Generalizing from the results of these studies, other methods of manipulating anxiety level should similarly affect the speed of learning complex material, i.e., lists with high inter-item competition.

Meprobamate has been shown to decrease hyperexcitability, anxiety, and aggression (1, 7, 14). It has also been shown to be of value in treating neurotic anxiety states (5, 7) and stress reactions associated with tension headaches (3). If meprobamate also decreases the non-specific anxiety or emotionality of the type measured by the Taylor Scale, it should lead to similar effects in learning a competitive list of items. The prediction would be as follows: Individuals under meprobamate will perform better in learning a competitive list of items than individuals under placebo. This study is a test of this prediction.

B. METHOD

1. *Subjects*

Seventy subjects were selected from a pool of volunteers at the University of Michigan and were paid for participating. Nine subjects did not participate, thereby reducing the number of subjects to 61. All subjects were between 22 and 28 years of age. Both male and female subjects were employed. Subjects were randomly assigned to the meprobamate ($n = 29$) or placebo treatment ($n = 32$) and the drugs were administered in a double-blind fashion.

2. *Procedure*

The subjects were run in groups varying in size from 8 to 12 persons containing approximately an equal number of meprobamate and placebo subjects in each group. One and a quarter hours after administration of three times the usual dose of meprobamate (1,200 mg., a large dose employed to assure that any positive effects would be clearly apparent) or placebo (three similar capsules) the instructions were read and the subjects were given five trials on a practice list to acquaint them with the procedure. The pills containing meprobamate or placebo were identical in appearance. The test list was presented immediatly after the practice list. The total time between ingestion of the pills and the beginning of the test was one hour and 40 minutes. To increase comparability between the complex learning experiments of Spence, *et al.* (11, 12) on manifest anxiety and paired-associates learning, the same list (see Table 1) was employed as well as the same anticipation interval. This list consisted of 12 paired associates taken from Haagen's word list (4). These word pairs had been chosen so as to produce high inter-item competition. The procedure of this experiment differs somewhat from that customarily employed in rote learning studies in that our subjects were run in groups. This procedure was thoroughly pretested and was found to be efficiently suited for learning studies involving drug ingestion followed by long waiting periods before testing. On any given word pair, a buzzer followed 1.7 seconds after the presentation of the stimulus word. Three seconds after the buzzer sounded, the response word of the pair was given. The learning task consisted of correctly anticipating and writing the response word of each pair before the buzzer sounded. Subjects did not have to complete writing the word before the buzzer in order that it be scored correct. It was only necessary that they begin writing the

TABLE 1
LIST OF TEST WORDS

Stimulus	Response
Barren	Fruitless
Arid	Grouchy
Desert	Leading
Little	Minute
Petite	Yonder
Undersized	Wholesome
Roving	Nomad
Gypsy	Opaque
Migrant	Agile
Tranquil	Placid
Quiet	Double
Serene	Headstrong

correct response during the 1.7-second interval and complete the writing of the word before the correct word was given. If the subject had not begun to write the anticipated word before the buzzer sounded, he was not to write at all. The group sat together in a room and an auditory presentation of the list was made. Although group testing is more efficient than individual testing, it is probably less sensitive since it is not possible to observe each subject individually as to his understanding and obeying of instructions during the test period.

C. Results and Discussion

The major findings are shown in Figure 1. The rate of learning the list is higher for the meprobamate subjects than for the placebo subjects. Employing the learning criterion of two successive correct trials, the meprobamate subjects took a mean of 13.7 trials to reach the criterion, while the placebo group required 16.5 trials to reach criterion. This difference in

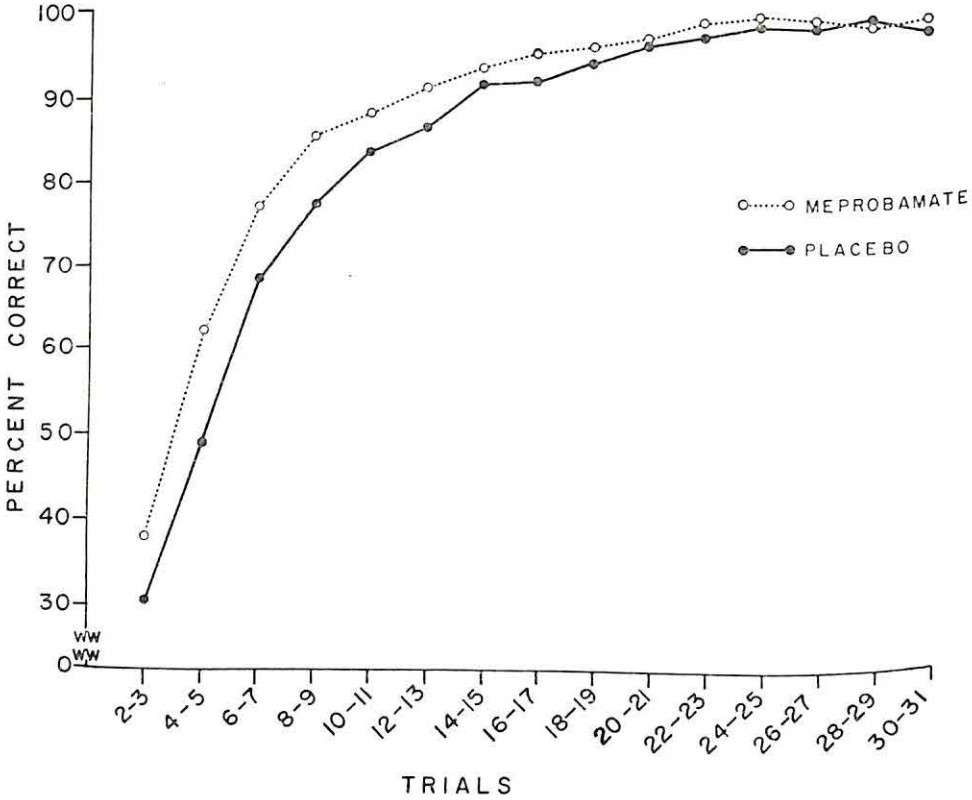


FIGURE 1

Paired-Associates Learning as a Function of Meprobamate vs. Placebo Ingestion the speed of learning is in the predicted direction and is significant ($t = 1.76$, $p < .05$ by a one-tailed test). This finding is suggestive from three points of view. First, to the learning theorist it points to an operation by which drive level may be manipulated experimentally in situations where anxiety is aroused by the organism's perception of the task rather than by a fairly well-controlled stressor like electric shock. Second, to the psychopharmacologist it provides additional information as to the drive-altering properties of meprobamate. Finally, to the psychiatrist or applied psychologist the findings suggest that when there is a high degree of competition among a group of items to be learned, administration of meprobamate may increase the rate at which the items are learned.

D. SUMMARY

The present experiment compared the effect of meprobamate vs. placebo on paired-associates learning involving a list with high inter-item competition. It was found that subjects under meprobamate learned the material significantly more rapidly than subjects under placebo. Group testing was employed. This outcome was predicted from Hull's conceptualization of anxiety or emotionality as directly reflecting the drive level.

REFERENCES

1. BERGER, F. M., HENDLEY, C. D., & TYNES, T. E. Effect of meprobamate (Miltown) on animal behavior. *Federation Proceedings*, 1956, **15**, 400 (Abstract).
2. FARBER, I. E., & SPENCE, K. W. Complex learning and conditioning as a function of anxiety. *J. Exp. Psychol.*, 1953, **45**, 120-125.
3. FRIEDMAN, A. P. The treatment of chronic headache with meprobamate. *Ann. N. Y. Acad. Sci.*, 1957, **67**, 822-828.
4. HAAGEN, C. H. Synonymity, vividness, familiarity, and association value ratings of 400 pairs of common adjectives. *J. of Psychol.*, 1949, **27**, 453-463.
5. HOLLISTER, L. E., ELKINS, H., HILER, E., & ST. PIERRE, R. Meprobamate in chronic psychiatric patients. *Ann. N. Y. Acad. Sci.*, 1957, **67**, 789-801.
6. MONTAGUE, E. K. The rôle of anxiety in serial rote learning. *J. Exp. Psychol.*, 1953, **45**, 88-91.
7. OSINSKI, W. A. Treatment of anxiety states with meprobamate. *Ann. N. Y. Acad. Sci.*, 1957, **67**, 766-772.
8. PHILLIPS, R. E. Use of meprobamate (Miltown) for the treatment of emotional disorders. *Amer. Practit. & Dig. Treat.*, 1956, **7**, 1573-1576.
9. SPENCE, K. W. Behavior Theory and Conditioning. New Haven: Yale Univ. Press, 1956.
10. SPENCE, K. A theory of emotionally based drive (D) and its relation to performance in simple learning situations. *Amer. Psychol.*, 1958, **13**, 131-141.
11. SPENCE, K. W., FARBER, I. E., & MCFANN, H. H. The relation of anxiety (drive) level to performance in competitional and non-competitional paired associates learning. *J. Exp. Psychol.*, 1956, **52**, 296-305.
12. SPENCE, K. W., TAYLOR, J., & KETCHEL, R. Anxiety (drive) level and degree of competition in paired associates learning. *J. Exp. Psychol.*, 1956, **52**, 306-310.

13. Taylor, J. A personality scale of manifest anxiety. *J. Abn. & Soc. Psychol.*, 1953, **48**, 285-290.

14. Yen, H. C. Y., Stranger, R. L., & Milliman, N. Isolation-induced aggressive behavior in ataractic tests. Paper read at Abel Centennial Meeting, American Society for Pharmacology and Experimental Therapeutics. Baltimore, Sept., 1957 (Abstract).

Department of Psychology
University of Michigan
Ann Arbor, Michigan

Published as a separate and in *The Journal of Psychology*, 1959, **47**, 87-98.

AN EXPERIMENT ON THE MEDIATION OF TRANSFER IN PAIRED-ASSOCIATE LEARNING*

Department of Psychology, Washington University

JOHN L. CRAWFORD AND JAMES M. VANDERPLAS

A. INTRODUCTION

The concept of mediation[1] has been employed by many investigators to account for results of experiments on transfer of training and other phenomena associated with the learning of verbal material. Early theorists, such as Titchener (23) and Wundt (26), employed the basic ideas of mediation in discussions of effects of associations, made in the past experiences of their observers, upon the formation of new associations. More recently, theorists such as Goss (7), Hebb (9), Hull (10), Mowrer (14), and Osgood (15) have elaborated the mediation hypothesis to account for a variety of phenomena, ranging from the development of simple phase-sequence relations in perception (Hebb, 9) to the development of meaningfulness of complex signs (Osgood, 15, Ch. 16) and changes in the significance of object stimuli (Mowrer, 14).

Specifically relevant to the present paper are the accounts of the development and transfer of verbal associations stemming from the mediation hypothesis. A number of investigators have employed mediation to account for certain transfer effects in verbal learning, involving the establishment of prior S-R connections and employment of them in learning new materials. Other theorists have employed mediation to account for early transfer in the development of meaning (15), language behavior (4, 5, 17, 18), learning (7, 14), and social development (9, 13).

But while the concept of mediation, by itself, is fairly clear as to its logical reference, the variables associated with it have more often been inferred than experimentally varied. Relatively few attempts, largely of a demonstrational character, have been made to test experimentally the various associated hypotheses concerning mediation. The studies by Birge (1),

---

*Received in the Editorial Office on October 28, 1958, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] Mediation is used in this paper to refer to an intervening process involving the generation of cue-stimuli by reactions to object-stimuli. In this sense, the mediation hypothesis referred to throughout this paper is analogous to the hypotheses made by theorists such as Hull, concerning what he called the "pure stimulus act" (10, p. 515) and to the same hypothesis as elaborated by Osgood (15, pp. 392-412).

Bugelski (2), Peters (16), and Shipley (20) are representative of attempts to establish mediating associations experimentally. The studies by Buzzotta (3), Osgood (15), and Russell and Storms (19) represent attempts to demonstrate the effects of previously established associations upon the learning of new associations.

Many of the studies employing the hypothesis of mediation may be subjected to one or both of two common criticisms: either (*a*) they were designed primarily as demonstrational experiments and thus were not designed to determine the functional relations involved in mediation or (*b*) they were not designed to include sufficient controls to rule out possible effects of alternative variables or processes. Other effects, such as transfer on the basis of general factors, e.g., learning sets (8), warm-up (22) and the like, and specific factors, e.g., positive and negative transfer effects (11), may also affect the acquisition of new verbal associations. But little attention has been paid to the possible relation between these alternative accounts of transfer and the mediation hypothesis.

B. THE EXPERIMENT

1. *Purpose*

The present study was conducted with two purposes in mind: The first was to test for effects of mediation when the mediating associations are established experimentally and prior to the task in which transfer occurs. In the study by Buzzotta (3), for example, the mediating associations were established, not experimentally within the framework of the study, but rather on the basis of selection of words which arouse existing associations within a framework of common language habits. While this is not undesirable in itself, it is not often known precisely the extent to which the mediation is a result of possibly additional variables which contribute to the common language habits. If the learning of verbal material were facilitated by mediating associations established within the experiment itself, there would be more cogent evidence in support of the mediation hypothesis.

The second purpose of the study was to provide in the experimental design for control of effects of alternative factors such as those mentioned above. In the studies surveyed, at least part of the facilitation of learning could be accounted for on the basis of factors other than mediation *per se*. That is, general practice effects, learning sets, warm-up, and specific transfer effects (both positive and negative) may account for part of the transfer effect.

To accomplish the first of these two purposes, verbal associations were established prior to the transfer tasks, but within the same experiment. For the experimental subjects three lists of nonsense syllables, of the form B-C, A-B, A-C, were used. Learning of the B-C list was assumed to establish associations which would facilitate the learning of the A-C list following learning of the A-B list. To accomplish the second purpose, control groups learned lists under conditions designed to include alternative effects noted above: three control groups learned lists of the form B-C, D-E, A-C; D-E, A-B, A-C; and D-E, F-G, A-C, respectively. These groups served to provide for alternative effects occurring when three lists are learned. Three additional control groups served to provide for alternative effects occurring when two lists only are learned. These three additional groups learned lists of the form B-C, —, A-C; —, A-B, A-C; and D-E, —, A-C, respectively. Finally, a basic control group learned only one list, A-C. The complete paradigm of the experiment is shown in Table 1. The control groups C-1 to C-7 represent an exhaustive set within the restrictions of the paradigm applied to the *E* group. That is, both positive and negative transfer effects (C-1 and C-4; C-2 and C-5, respectively), as well as general practice effects and warm-up (C-3 and C-6) are represented.

TABLE 1
EXPERIMENTAL DESIGN PARADIGM

Group	Condition		
E	B-C	A-B	A-C
C-1	B-C	D-E	A-C
C-2	D-E	A-B	A-C
C-3	D-E	F-G	A-C
C-4	B-C	—	A-C
C-5	—	A-B	A-C
C-6	D-E	—	A-C
C-7	—	—	A-C

If the mediating associations established during learning of the B-C list facilitate transfer to the A-C list to an extent greater than that afforded by the control conditions, then one would expect that the *E* group would learn the A-C list in significantly fewer trials than does any one of the control groups.

2. *Subjects*

Eighty airmen, students in a course for airborne radio mechanics, at Scott Air Force Base, served as subjects in the experiment. Ten men were assigned at random to each group, one experimental and seven control groups, as indicated in Table 1.

3. *Materials and Apparatus*

A set of five lists, each containing eight pairs of nonsense syllables, was constructed from the 53- and 60-per cent association value syllables of Glaze's lists (6). Each list was arranged in four different sequences to avoid serial effects. Selection and randomization of the lists were carried out in accordance with the criteria set forth by Hilgard (21, p. 540).

A Lafayette memory drum was used to present the lists. Each stimulus syllable was presented alone for about 2 sec., followed by the stimulus and response syllables together for about 2 sec., with a pause of about four seconds between trials.

4. *Procedure*

The subjects entered the darkened experimental room individually. The subject was seated at the illuminated memory drum and asked his name, level of education, and age. He was then instructed in the usual manner, to anticipate each response syllable when the stimulus syllable appeared and to associate each S-R pair with each other. The criterion of mastery for all lists was three correct anticipations of all eight items. A rest pause of five minutes was employed between the learning of each list. All responses were recorded.

C. RESULTS

Means and *SD*'s of trials to the criterion of mastery of all the lists are shown in Table 2, for all groups, experimental and control. A test of homo-

TABLE 2

MEANS AND STANDARD DEVIATIONS FOR EXPERIMENTAL AND CONTROL GROUPS

Group	*M*	*SD*	*M*	*SD*	*M*	*SD*
E	B-C 39.6	13.61	A-B 39.2	19.75	A-C 27.8	11.46
C-1	B-C 48.7	16.13	D-E 41.8	11.35	A-C 33.0	13.68
C-2	D-E 34.1	14.22	A-B 39.6	20.27	A-C 33.3	15.09
C-3	D-E 45.9	11.79	F-G 41.5	16.93	A-C 28.1	12.42
C-4	B-C 44.1	14.76			A-C 26.0	6.12
C-5			A-B 39.5	9.95	A-C 26.6	7.74
C-6	D-E 41.9	13.43			A-C 31.6	14.19
C-7					A-C 41.9	14.94

geneity of variance of the data for the A-C lists [Bartlett's test (24)] yielded a value of B equaled to 10.81 which, when referred to a Chi-square table was found to be insignificant at the 10 per cent level (Chi-square .90 for 7 degrees of freedom is 12.00). The hypothesis was therefore accepted that the variances in trials to mastery are equal for all eight groups.

TABLE 3
ANALYSIS OF VARIANCE OF THE A-C LISTS

Source of variance	Sum of squares .	df	Variance estimate	F
Between grps.	1,914.59	7	273.51	1.579
Within grps.	12,464.3	72	173.12	
Total	14,378.89	79		

Analysis of variance of the trials data for the A-C list is summarized in Table 3. Comparison of the Between Groups mean square (273.51) with the Within Groups mean square (173.12) yielded an F-ratio of 1.579 which, for 7 and 72 degrees of freedom is insignificant at the 5 per cent level. On the basis of this analysis it may not be concluded that the various groups differ in respect to the mean number of trials to mastery.

Returning briefly to Table 2, it may be seen that the basic control group (C-7) required a mean of 41.9 trials to master the A-C list, while the experimental and remaining control groups required between 26.0 (C-4) and 33.3 (C-2). In an analysis of the type summarized in Table 3, it is possible to obtain an F-ratio which is not significant even though significant differences exist between some of the individual pairs of means (25). The hypothesis may be advanced that even though the experimental group does not differ from the control groups which include alternative effects, all groups (other than the basic control group) may differ from the basic control which does not include any effect, including mediation, except practice on the A-C list. If supported, this hypothesis would indicate that positive transfer occurred in the control groups as well as in the experimental group and that transfer may occur as a result of factors other than mediation (but not in any mutually exclusive way). Accordingly, a test was made of the hypothesis that the basic control group (C-7) and the combined experimental and Groups E and C-1 to C-6 do not differ. A t-test was employed and is summarized in Table 4. The obtained t (2.84) is significant beyond the one per cent levl, indicating that the combined groups do show some joint effects of specific or general factors or mediation which result in significant positive transfer.[2]

---

[2] An additional t-test comparing the highest mean (33.3) with the lowest mean

TABLE 4
SUMMARY OF t-TEST COMPARING BASIC CONTROL GROUP C-7 WITH COMBINED EXPERIMENTAL (E) AND CONTROL GROUPS C-1 TO C-6

Group	Mean	SE_{md}	t	p
C-7	41.9	4.364	2.84	$<.01$
Combined E and C-1 to C-6	29.5			

There are several hypotheses which may account for the lack of significance of differences among the means of the experimental and control groups other than the basic control. First, it may be that the groups themselves were so constructed or selected that an interaction occurred between the groups and conditions, thus obscuring the mediation effect. Second, differential difficulty of the first and second lists learned may also have interacted with other effects to cloud differences. Third, combined effects of specific positive transfer (e.g., B-C, A-C paradigm from first to third list in the E group) or negative transfer (e.g., A-B, A-C paradigm from second to third list) may have contributed to the lack of clear differences. Accordingly, three analyses of covariance were performed on the data. The first was conducted to control for effects of first-list learning in groups which learned three lists; the second was conducted to control for effects of both first- and second-list learning in these same groups. The third was conducted to control for effects of first-list learning in groups which learned only two lists. These analyses would serve not only to (a) increase the precision of the experiment, but also to (b) permit a comparison of the data for the A-C lists with the same data after an arithmetic correction had been applied to account for the effects of group differences, first- and second-list learning, etc. One might expect that all groups would approach the performance of the control group when their scores were corrected for the predicted value in a covariance design. Although the covariance adjustment method does not *necessarily* reduce the differences between the means (12), a comparison of the actual data with the predicted data would afford a partial test of hypotheses concerning the effectiveness of the various controls.

The first analysis of covariance of the data for the experimental group and control Groups C-1 to C-3 is summarized in Table 5. The results of this analysis indicate that the precision of the experiment has not been increased appreciably by the control of first-list learning scores in these

---

(26.0) in Table 2 did not reveal a significant difference between the most divergent control groups. Thus, while the combined control groups differ from the basic control group, they do not differ among themselves, even when a biased selection of means is made (12) to permit the difference to appear.

TABLE 5
SUMMARY OF ANALYSIS OF COVARIANCE FOR GROUPS *E* AND C-1 TO C-3, CONTROL FOR FIRST LIST

Source	*df*	Sum of squares	Variance estimate
Between groups	3	637.00	212.30
Within groups	35	4,781.54	136.61
Total	38	5,418.54	
Group regression	3	157.76	52.55
Within groups residual	32	4,623.78	144.50
Group mean residual	2	266.00	133.00
Combined residual	1	371.00	371.00

groups. Comparison of the Between Groups variance (212.3) with the Within Groups variance (136.61) yielded an *F* of 1.54 which, for 3 and 35 degrees of freedom is not significant at the 5 per cent level. Additional tests, of linearity of regression and common slope as well as a test of zero slope all were consistent with the null hypothesis. Thus, the amount of the decrease in means from the first list to the third does not vary significantly around a common regression line, computed from the scores for the first. If any differences do exist between means of the scores for the A-C lists, they may not be attributed to effects of differences between means of scores on the first list (e.g., group differences in learning the first list, differential list difficulty, etc.).

The second analysis of covariance of the data for the experimental group and control Groups C-1 to C-3 with control now for both first- and second-list scores, is summarized in Table 6. This analysis indicates that the precision of the experiment has been increased by the additional control. Comparison of the Between Groups variance (955.00) with the Within Groups variance (97.24) yielded an *F* of 9.82 which, for 3 and 35 degrees of freedom is significant beyond the 0.1 per cent level. The hypothesis of common slope may thus be rejected, leading to the conclusion that significant variation exists among the groups in scores for the A-C list. However, the control in the analysis was a control for effects of both first- and second-list scores. We may infer from the experimental design that the effects of differential performance on these lists have been controlled and that the resulting residual variance, around the regression plane, reflects variation in performance in the various subject groups. Differential effects which occurred during learning of the first and second lists may have contributed to the lack of significant differences found in the analysis of Table 3. If so, then there may also be an interactive effect, of both first and second

TABLE 6
SUMMARY OF ANALYSIS OF COVARIANCE FOR GROUPS *E* AND C-1 TO C-3, CONTROL FOR FIRST AND SECOND LISTS

Source	*df*	Sum of squares	Variance estimate
Between groups	3	2,865.00	955.00
Within groups	34	3,306.30	97.24
Total	37	6,171.30	
Groups regression	6	30.72	5.12
Within Groups residual	28	3,275.58	116.98
Group mean residual	1	209.39	209.39
Combined residual	2	2,655.60	1,327.80

lists jointly, which obscures the group differences. When this joint effect is controlled, group differences are revealed.

A test for slope of the regression line based on second-list scores led to acceptance of the null hypothesis. Thus, very little of the variation of scores for the A-C list is accounted for on the basis of second-list scores alone. Additional tests for linearity of regression and homogeneity also permit acceptance of the null hypothesis. Thus, it appears that the differential performance of the various groups is not accountable on the basis of regression of the scores on the first- and second-list scores and that the variation is largely attributable to group differences.

The third analysis of covariance of the data for the control Groups C-4 to C-6 is summarized in Table 7. The analysis indicates an appreciable increase in the precision of the experiment as a result of control for first-list scores in these groups. Comparison of the Between Groups variance (2,798.55) with the Within Groups variance (90.06) yielded an *F* of 30.9 which, for 3 and 25 degrees of freedom is significant beyond the 0.1 per cent level. The hypothesis of common regression may be rejected, leading

TABLE 7
SUMMARY OF ANALYSIS OF COVARIANCE FOR GROUPS C-4 TO C-6, CONTROL FOR FIRST LIST

Source	*df*	Sum of squares	Variance estimate
Between groups	3	8,395.59	2,798.55
Within groups	25	2,251.56	90.06
Total	28	10,647.15	
Groups regression	2	1,404.18	702.09
Within groups residual	24	847.38	31.14
Group mean residual	1	188.00	188.00
Combined residual	1	8,207.59	8,207.59

again to the conclusion that significant variation exists among the various groups in scores on the A-C list. Here again, however, the differences are largely group differences, since differential first-list performance has been controlled.

There appears to be an effect on the A-C list scores which is due to first-list performance. The hypothesis of common slope was tested, and the resulting *F*, from comparison of the Groups Regression variance (702.09) with the Within Groups variance (31.14), was 22.6 which, for 2 and 24 degrees of freedom is significant beyond the 0.1 per cent level. Performance on the A-C list thus appears partly accountable on the basis of first-list performance, depending, perhaps, on the type of list learned first or differential practice among the groups. The additional hypothesis, of linearity of regression, was acceptable at the 5 per cent level.

One final analysis was conducted, to determine if, after arithmetic adjustment, the means of the experimental group and control Groups C-1 to C-6 might approximate the mean of the basic control group, C-7. This analysis is summarized in Table 8. In the analysis, it was necessary to compute regression coefficients, *b*, for total covariance, between groups covariance, and within groups covariance, separately for Groups *E* and C-1 to C-3, who learned three lists, and Groups C-4 to C-6, who learned two lists. Means for scores on the A-C list were then predicted from the scores on control lists, using appropriate regression coefficients as indicated in the

TABLE 8

REGRESSION COEFFICIENTS AND PREDICTED AND ADJUSTED MEANS FOR GROUPS *E* AND C-1 TO C-3 WITH FIRST- AND SECOND-LIST CONTROLS AND FOR GROUPS C-4 TO C-6 WITH FIRST-LIST CONTROLS

Regression coefficient		Group	Predicted mean	Adjusted mean (b_w)	(b_b)
b_1T	.1394	E	29.79	29.6	30.1
b_2T	.3271	C-1	31.89	28.6	33.1
b_1w	.6670	C-2	29.16	38.7	31.2
b_2w	.12	C-3	31.9	22.8	27.4
b_1b	—.59				
b_2b	2.9				
b_T	.63	C-4	29.51	24.53	26.23
b_w	.64	C-5	26.61	28.07	26.4
b_b	—.10	C-6	28.12	31.54	31.61

table, and adjusted according to the covariance method (24).[3] An inspection of the data in Table 8 indicates that the adjustment of the A-C means for the *E* and C-1 to C-3 groups, using the within groups coefficient, has accentuated the differences among the means, compared to the observed scores of Table 2. However, when the between groups coefficient is used, the differences are decreased relative to the observed values. A similar result appears for Groups C-4 to C-6. These results would seem to provide further support for the hypothesis that the major source of variance in the A-C list means is one involving group differences. It is also to be noted that the adjustment of the A-C list means did not result in their becoming closer in value to that of the basic control group.

D. Discussion

These data appear to provide little support for the mediation hypothesis, using the B-C, A-B, A-C paradigm for the experimental group. The analysis of variance of the A-C list data yielded an insignificant *F* when employed to test for differences in performance among the various groups. A *t*-test, comparing the combined experimental and control Groups C-1 to C-6 with the basic control, C-7, did support the hypothesis concerning possible alternative effects of general and specific transfer (and perhaps mediation) which may have resulted in significant positive transfer in all groups except the basic control. Nonetheless, the hypothesis that mediation provides for positive transfer over and above that provided by other, general and specific, factors was not supported.

The analyses of covariance indicated that the experiment was sensitive enough to reveal group differences when effects of previous lists were controlled statistically, and the covariance adjustment of the A-C list means appeared to support this notion further. However, the adjusted values of the A-C means did not approach the value for the basic control group. This latter finding would seem to indicate that some of the general practice effect or other contributor to the positive transfer obtained in all groups has not been accounted for by the design or the analysis. Nonetheless, the over-all effect of prior practice seems to be one largely accountable on the basis of positive transfer due to general practice with some transfer due to factors perhaps other than specific mediational factors inherent in the design.

---

[3] The adjustment is carried out numerically by subtracting from each of the means the gain expected on the basis of its association with the control list mean score above the grand mean or adding the loss expected on the basis of association with the control list mean score below the grand mean. The adjustment is *b* (*Group mean—Grand mean*), and the means are those of the control list scores. For the *E* group and control Groups C-1 to C-3, two adjustments are made, one for each of the first two lists. For control Groups C-4 to C-6, only one is needed.

These remarks need perhaps to be qualified somewhat, in that they are not *contrary* to the mediation hypothesis *as such*. Other investigators, using both the A-B, B-C, A-C paradigm, as well as the B-C, A-B, A-C paradigm, have found results in support of the hypothesis. Bugelski and Scharlock (2), for example, obtained positive transfer to an A-C list following learning of two prior lists, A-B, B-C. While the mean number of trials to criterion did not differ significantly for the various groups, the "mediation" group in their experiment was found to reach earlier stages of mastery (the first correct response) in fewer trials. However, these results could be alternatively interpreted as due to a combination of mediation, specific positive transfer, or associative interference in the control group. The possibility that alternative effects such as these might also account for the transfer obtained, tends to weaken somewhat the evidence for the clear-cut contribution due to mediation. It may also be true that mediation, as it is used in the present paper, while being effective as a means of producing transfer, is not more (or less) effective than other means such as those mentioned above. If so, then it would be difficult to isolate unless a fairly complex experimental design were used. The present findings tend to indicate that mediation alone contributes very little additional transfer over and above that contributed by general and specific factors. When these alternatives are controlled, as in the present study, differences between groups are either (*a*) small or (*b*) largely group differences.

REFERENCES

1. BIRGE, J. S. Verbal responses in transfer. Unpublished doctor's dissertation, Yale University, 1941.
2. BUGELSKI, B. R., & SCHARLOCK, D. P. An experimental demonstration of unconscious mediated association. *J. Exp. Psychol.*, 1952, **44**, 334-338.
3. BUZZOTTA, V. R. Association-probability in the study of verbal behavior. Unpublished doctor's dissertation, Washington University, 1956.
4. COFER, C. N., & FOLEY, J. P. Mediated generalization and the interpretation of verbal behavior: I. Prolegomena. *Psychol. Rev.*, 1942, **49**, 513-540.
5. ———. Mediated generalization and the interpretation of verbal behavior: II. Experimental study of certain homophones and synonym gradients. *J. Exp. Psychol.*, 1943, **32**, 168-175.
6. GLAZE, J. A. The associative value of nonsense syllables. *J. Genet. Psychol.*, 1928, **35**, 255-267.
7. GOSS, A. E. A stimulus-response analysis of the interaction of cue-producing and instrumental responses. *Psychol. Rev.*, 1955, **62**, 20-31.
8. HARLOW, H. F. The formation learning sets. *Psychol. Rev.*, 1949, **56**, 51-65.
9. HEBB, D. O. The Organization of Behavior. New York: Wiley, 1949.

10. HULL, C. L. Principles of Behavior (Ch. XII). New York: Appleton-Century, 1943.
11. McGEOCH, J. A., & IRION, A. L. The Psychology of Human Learning. New York: Longmans, 1952.
12. McNEMAR, Q. Psychological Statistics (2d ed.). New York: Wiley, 1955.
13. MILLER, N. E., & DOLLARD, J. Social Learning and Imitation. New Haven: Yale Univ. Press, 1941.
14. MOWRER, O. H. The psychologist looks at language. *Amer. Psychol.*, 1954, **9**, 660-694.
15. OSGOOD, C. E. Method and Theory in Experimental Psychology. New York: Oxford Univ. Press, 1953.
16. PETERS, H. N. Mediate association. *J. Exp. Psychol.*, 1935, **18**, 20-48.
17. RAZRAN, G. H. S. A quantitative study of meaning by a conditioned salivary technique. *Science*, 1939, **90**, 89-90.
18. REISS, B. F. Semantic conditioning involving the galvanic skin reflex. *J. Exp. Psychol.*, 1940, **26**, 238-240.
19. RUSSELL, W. A., & STORMS, L. Implicit verbal chaining in paired-associate learning. *J. Exp. Psychol.*, 1955, **49**, 287-293.
20. SHIPLEY, W. C. An apparent transfer of conditioning. *J. Gen. Psychol.*, 1933, **8**, 382-390.
21. STEVENS, S. S. (*Ed.*). Handbook of Experimental Psychology. New York: Wiley, 1951.
22. THUNE, L. E. Warm-up effect as a function of level of practice in verbal learning. *Amer. Psychol.*, 1950, **5**, 251 (Abstract).
23. TITCHENER, E. B. A Textbook of Psychology. New York: Macmillan, 1919. (Pp. 386-389.)
24. WALKER, H. M., & LEV, J. Statistical Inference. New York: Holt, 1953.
25. WEBB, W. B., & LEMMON, V. W. A qualification in the use of analysis of variance. *Psychol. Bull.*, 1950, **47**, 130-136.
26. WUNDT, W. Lectures on Human and Animal Psychology. (Trans. by Titchener and Creighton.) New York: Macmillan, 1901.

Department of Psychology
Washington University
St. Louis 5, Missouri

Published as a separate and in *The Journal of Psychology*, 1959, **47**, 99-105.

FIELD WORK IN THE UNDERGRADUATE PSYCHOLOGY CURRICULUM*

Department of Psychology, New York University

ALICE GUSTAV

A. INTRODUCTION

Undergraduate students in psychology rarely have the opportunity to do more than become acquainted with subject matter through textbook readings and, for this reason, often miss the real impact of the material. If a student proceeds to graduate school in psychology he will, of course, obtain some type of practical experience such as an interneship. However, at Washington Square College in New York University we have found that many undergraduates enroll as psychology majors or minors with no intention of ultimately becoming psychologists. Students not oriented toward the profession tend to fall into three groups: (*a*) those who will enter graduate training in occupations dealing with people and who therefore desire a background knowledge of psychology, e.g., the ministry, medicine, social work, and teaching; (*b*) those who wish a general liberal arts education rather than pre-professional training; and (*c*) those who are yet in the process of seeking a vocational choice.

To make undergraduate psychology courses more meaningful to all students, and especially to help the last named groups, the author undertook responsibility for a program of supervised field work which students could do on a voluntary basis.

B. DESCRIPTION OF THE PROGRAM

In New York City there exist a large number and wide variety of community organizations which will supervise undergraduate volunteers. Through personal visits, correspondence, and telephone calls the author has prepared a list of such agencies. This is made available at the beginning of each term to students registered in advanced undergraduate psychology courses. If interested, they then arrange their own appointments for interviews and acceptance by the institutions involved. Each student is expected to work three hours per week at the agency selected. Approximately two weeks prior to the semester's end a request is sent to each organization asking that super-

---

*Received in the Editorial Office on November 4, 1958, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

visors fill out a form evaluating the volunteers. The evaluations are retained on file in the psychology department and are extremely useful when writing letters of recommendation for students. Often, graduate schools and prospective employers request statements about an applicant's ability to get along with people and his sense of responsibility, as well as academic grades. With these evaluations on hand it is possible to give evidence of actual performance rather than generalized statements based solely on classroom observation.

The present paper summarizes results of a follow-up study designed to learn what benefits students were deriving from this program. An anonymous questionnaire, with a covering letter, was mailed to the first 300 volunteers who had completed at least one semester of field work. The questionnaire called for information concerning: (*a*) agencies at which volunteers served, (*b*) type of work done, (*c*) reasons for electing to do field work, (*d*) whether the work was enjoyed, (*e*) whether it was a worthwhile experience, and (*f*) effect of field work on academic studies.

C. RESULTS

Of the 300 students, 119 responded (46 male and 73 female), or 39.7 per cent of the total. While it is not possible to determine whether the respondents constitute a representative sample of the original group, results for those who did answer were impressive quantitatively, and illuminating when examined qualitatively. Hence it appeared of value to present the findings for those people interested in techniques of enriching the undergraduate psychology curriculum.

Discussion in the balance of this paper will be restricted to the group of 119 respondents. Data for males and females have been combined as their answers to the questionnaire were remarkably similar.

1. *Agencies*

The 119 respondents worked in 57 different agencies. These included 33 settlement houses and community centers, 17 branches of the YM-YWCA and YM-YWHA, the Jewish Board of Guardians, three hospitals, and the New York City Child Guidance Bureau.

2. *Types of Work*

The students worked primarily as aides to leaders in game rooms, teen lounges, play groups, club meetings, arts and crafts groups, athletics, and classrooms of various kinds. Sixteen students escorted disturbed children to and from psychiatric appointments. A few performed duties such as editing a newspaper, assisting in program planning, and observing and recording the behavior of a specific child.

Mere enumeration of types of positions held cannot do justice to the variety of experience obtained. In one institution being an assistant in a game room might mean custodial work, i.e., preventing fights among the youngsters, handing out and retrieving play material, etc. In another agency it would involve considerable individual and personal attention to each child. In some places the student might deal with different children weekly, while in others they might find a closely knit group who return as a unit at each meeting.

3. *Reasons for Election of Field Work*

In response to the question "Why did you decide to do field work?" many students gave more than one reason. Therefore, the items described below total more than 119. The answers broke down into several main categories. Eighty-five desired the practical experience. A typical statement was "I felt that no written work could ever give me as much experience and knowledge about adolescents as actually working and being with them." Twenty-one students felt this would be an opportunity to determine whether they liked working with people as a vocation, e.g., "I was contemplating teaching as a profession and thought this would give me a good chance to decide whether I was temperamentally suited for the work." Thirteen were intrigued by what sounded like a new adventure, e.g., "Presented the prospects of an interesting and unusual experience not ordinarily found in undergraduate classes." Fifteen had already done such work, liked it, and were eager to continue. Eleven signified interest in working with people. Three wished to do it as an expression of altruism.

4. *Enjoyment of Work*

In response to the question "Did you enjoy the work? Why?" 105 said yes, six said partially, and eight said no. Again, some students gave more than one reason.

Of those respondents who answered in the affirmative, 31 said they had enjoyed it because they had learned new things, 25 felt they were contributing to others, 15 thought it related actualities to classroom material, 14 found it challenging and stimulating, 10 enjoyed working with children or adolescents, 6 "had fun," 5 learned about themselves, 3 found it active rather than passive, and 12 gave no reasons. The following were typical responses. "I was really learning three things: How you deal with adolescents in trying to teach them, the techniques in arts and crafts, and how they and I reacted to each other." "Offered opportunity to study and learn from a group brought up in totally different environmental surroundings. " I felt I was performing a worthwhile community service."

An interesting point is that 25 reported enjoyment because it gave them satisfaction to contribute to the welfare of others. Yet, as noted in the previous section, only three students had verbalized altruistic reasons as the motivating force in their decision to do volunteer field work. This is part of a life philosophy which parents and teachers so often attempt to inculcate in young people. Ordinarily, it is difficult because adolescents in our culture tend to hear such aims preached but are given little chance to discover for themselves the profound gratification to be derived from work of this nature. Apparently, volunteer field work accomplished this aim for some students.

Of those respondents who indicated they had not enjoyed the work, or that it had been just partially enjoyable, four felt they were only "policemen" or "glorified baby sitters," three felt inadequately prepared, two felt there was insufficient supervision, one did not think the returns were sufficient, one simply ran a movie projector, one found it boring and lacking in challenge, one said the facilities were not adequate to do a worthwhile job, and one gave no reason.

In view of the fact that an overwhelming majority of respondents were enthusiastic, it appears likely that the few who did not find it enjoyable had personal idiosyncracies which made them react in this way. For example, three students reported they did not enjoy the situation because they felt inadequately prepared to cope with it. Yet, when explaining the opportunity to students at the beginning of each term the author always carefully emphasizes that this is meant to be a first work experience and is definitely not intended for trained individuals. Therefore, students start their duties knowing we anticipate they will improve on the job, but that they are novices and not expected to be competently trained in advance. Also, two students complained of insufficient supervision and one found the lack of adequate facilities a hindrance to enjoyment of the work. It is all too true that a shortage of paid personnel and lack of equipment is a realistic problem in most community agencies. Nevertheless, as will be noted in the next section, other students found this a challenge to their own versatility.

5. *Value of the Experience*

When asked "Did you feel it was a worthwhile experience? Why?" 107 said yes, two said partially, and 10 said no. The eight who had replied they did not enjoy it were among the 10 who did not consider it worthwhile.

Again, some students gave more than one reason and therefore the total below is greater than 119.

Of those respondents who answered affirmatively, 56 indicated they had obtained a better understanding of children and adolescents, 28 felt it had supplemented the theory gained in class, 28 said their own feelings of responsibility, accomplishment or patience had been bolstered, 17 found it of value in relation to their future work, 11 considered it worthwhile because they learned about agencies and types of work, three learned to observe carefully, and four gave no reasons. Typical of their answers were the following. "It helped me to realize the importance of environmental influences upon children's lives—and the need for teachers and leaders who understand this fact. It also influenced my choice of a major in graduate school of educational psychology." "From my own point of view, i.e.—for future reference in my life, it helped me to finally decide against becoming a teacher, so that I directed my energies toward a profession for which I am better suited intellectually and psychologically. At the same time, I can see the value this experience would have for future teachers and social workers, etc." "Working with younger children gives one a chance to assume adult responsibilities, and to learn how to handle a variety of situations and to develop one's creativity as is called for in these activities." "Prepare me for being a parent." "We learned how to acclimate ourselves to working with the bare minimum of equipment and still do a creditable job."

An important finding in the above data is that 28 students reported field work had bolstered their own feelings of responsibility, accomplishment, or patience. As in the case of altruism, these are traits which adults would dearly love to see displayed by adolescents. It is a common complaint among the elders that such characteristics are woefully lacking in young people. Apparently what is required for many teen-agers is to be placed in a position where it is evident that being on time, following through on a task, etc., is truly important. Preaching by parents and teachers is usually of little avail, and it is only when an individual assumes such an assignment for the first time that the neeed for these qualities becomes obvious to him.

Another striking point is that many of the students, among the 11 who said they had learned about agencies and types of work and the 17 who found field work important in relation to their future careers, were given further clues in deciding about their own vocations. The sample comments above contain examples of a positive and a negative decision. Students often have unrealistic views about occupations. Field work can be one step in aiding an individual to ascertain, partially at least, whether he likes working with and for people, or whether he prefers a more solitary work environment.

Some students find they are more comfortable with one age group rather than another, e.g., a student considering early elementary school teaching as a profession may discover teen-agers more congenial than kindergarten children. Certainly it is better to learn this prior to entering graduate school training.

Becoming acquainted with the organization and activities of community agencies, as noted by 11 students, can be important in several ways. The individuals may wish to prepare for work in such agencies, and it is wise to develop some amount of familiarity with them in advance. Members of their families may wish to make use of such community organizations and they can help in evaluating the appropriate one. Finally, as citizens and taxpayers in a community, they should be aware of what facilities are available and how they can be improved.

Of those respondents who replied they did not find the experience a worthwhile one, or indicated that it was only partially so, four did not believe it rewarding enough in relation to the amount of time put in, three were hampered by the lack of guidance, two would have liked different groups of children, one found the inadequate facilities frustrating, and two gave no reasons. A typical answer was "The realization of the unique problems of the individual, of the variation in the reaction of individuals to similar problems and of environmental variation within what seemed ostensibly the same environment was enlightening, but hardly rewarding in relation to time spent."

As in the previous section, it can be noted that an overwhelming majority did find it a worthwhile experience, and that the reasons given by a few for not regarding it useful were the very ones a greater number of students gave for calling it a challenging situation.

6. *Effect of Work on Academic Studies*

In response to the question "Did the work help in your academic studies? How?" 65 said yes, 10 said partially, and 44 said no. An analysis of the responses indicated that the students had interpreted this question in two different ways. Some, in saying yes, meant that it helped them directly on examination material. Others felt it was advantageous because it supplemented their understanding of real life situations without actually aiding them directly on class tests, e.g., "This work improves upon the textbook and the lecture. To read and listen can give a foundation, but only actual experience can give the student a true concept of the work and problems met on the outside." Liberal arts courses are apt to be on an abstract, theoretical level. Field work can be one attempt to show students how, by themselves, they can learn to apply such abstract knowledge to the life situations they will later meet.

Those who answered negatively almost unanimously agreed that it was because (*a*) the class examination material was not on the work done at the agencies or (*b*) their major was in a remote field, e.g., "My predominant interest has been in experimental animal psychology. This field work did not contribute to my ability to work in that area."

The prime reason for answering that it was partially useful in academic studies was that the students became aware, some time after the field work, that it had been useful and they had not taken sufficient advantage of it, e.g., "Not till I began studying for my Master's degree in education did I realize how beneficial this type of work was."

D. CONCLUSIONS

Field work which, in so many schools, is reserved for graduate level students can serve useful, though different, functions at the undergraduate level. At the latter stage it is not designed to develop competence in professional training. Rather, as indicated by the answers of the students in this report, we find the experience can provide young people with the opportunity to:

1. Learn, to some extent, whether they enjoy occupations which involve working with people.
2. Learn to develop a sense of responsibility in a job situation.
3. Derive a feeling of accomplishment through service to others.
4. Learn how to apply academic, theoretical training to real life situations.
5. Become acquainted with the organization and activities of community agencies.
6. Obtain a first supervised work assignment, which is generally difficult for undergraduate students, even on an unpaid basis.
7. Provide the psychology department with concrete data about the student's abilities and personal characteristics which can be utilized to make letters of recommendation truly meaningful.

Department of Psychology
Washington Square College
New York University
Washington Square
New York 3, New York

Published as a separate and in *The Journal of Psychology*, 1959, **47**, 107-116.

SHIFTS IN CONNOTATIVE MEANING OF WORDS AS A FUNCTION OF VARYING AMOUNTS OF PREVIOUS RESTRICTIVE EXPERIENCE*[1]

Bell Telephone Laboratories and New York University

M. S. MAYZNER AND M. E. TRESSELT[2]

A. PROBLEM

Previous work (2) has shown that words, scaled as to their frequency of inclusion in given conceptual categories, will be judged as belonging or not belonging to these categories as a function of the absolute scale value of the word and antecedent anchoring conditions. Thus, *S*s, given words having low scaled frequencies of inclusion in particular concepts, judged more such words as belonging to these concepts, than *S*s who had previously been anchored by judging words having high scaled frequencies of inclusion, before judging the low scaled words. The question was raised "What is the optimum number of discrete stimuli or words needed to produce the strongest anchoring effect?" The present study examines this issue, by systematically varying the number of anchoring stimuli.

B. METHOD

1. *Materials*

The stimulus material presented to *S*s consisted of a list of words. These words were obtained and scaled from previous studies (1, 3, 5) as follows: One hundred *S*s were presented with a list of 600 words and were asked to check each word as belonging to none, 1, 2, 3, 4, 5, or 6 concepts. The six conceptual categories were theoretical (*T*), economic (*E*), aesthetic (*A*), social (*S*), political (*P*), and religious (*R*).

By summing the number of checks for each word for the 100 *S*s, 600 word profiles were obtained (3), which showed the frequency with which each word was judged as belonging to each of the six concepts. For example, the word profile for the word "science" showed frequency counts of 90, 21, 9, 23, 20, and 7 for the *T, E, A, S, P,* and *R* concepts, respectively, indicating that 90 out of 100 *S*s checked "science" as belonging to the theoretical concept, 21 out of 100 *S*s checked it as belonging to the economic concept, etc.

---

*Received in the Editorial Office on November 6, 1958, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This study was supported in part by a grant made to M. E. Tresselt by the New York University Arts and Science Research Fund.

[2] The authors wish to express their thanks to Dr. Paul S. Olmstead of Bell Telephone Laboratories for his helpful advice on the statistical evaluations in this paper.

TABLE 1
THE WORDS COMPOSING LISTS *A* AND ***B***

List *A**					
Theoretical	Economic	Aesthetic	Social	Political	Religious
science	sale	design	conduct	assassinate	faith
reason	receipt	easel	sex	king	priest
evidence	industry	imaginative	people	senator	sin
verification	expenditure	poem	friend	vote	hymn
rational	depreciation	nature	carnival	socialism	circumsize
conclusion	distribution	blend	fraternity	ally	sermon
velocity	petroleum	horizon	reunion	impeach	spirit
synthesize	wages	music	team	patronage	ritual
method	price	soprano	pet	veto	Easter
hypothesis	cost	beauty	gathering	nominate	confession
analysis	rates	color	games	execute	blessing
evolution	trade	sonnet	hug	lobby	crucifix
validity	invest	paint	popular	nation	devil
postulate	budget	museum	farewell	representative	sect
empirical	assets	harmony	cabaret	elect	cathedral
experiment	coal	composition	everybody	administration	pew
prove	investment	draw	guest	campaign	salvation
inductive	monopoly	verse	meet	legislation	theology
dimension	imports	picture	picnic	force	temple
cyclotron	expense	creative	dear	amendment	orthodox
formula	auditor	classical	amiable	mayor	holy
laboratory	buy	statue	community	peace	prayer
research	merger	overture	wedding	rule	soul
solution	inventory	exhibit	bare	republic	reverend
phenomenon	finance	artist	brother	congress	pagan
elemental	account	literary	joke	court	Bible
equilibrium	market	opera	tea	run	divine
universe	purchase	sonata	vacation	election	heaven
infinite	money	shade	wife	dictator	piety
equation	ledger	ballet	family	constitution	miracle

*Words in List *A* were randomized for presentation to *S*.

These word profiles were used to construct two basic lists of words having properties as follows: List *A* (shown in Table 1) contains words having

TABLE 1*a*

LIST *B***

1. table
2. second
3. trunk
4. blanket
5. turn
6. last
7. substitute
8. blow
9. veil
10. ass
11. hand
12. entry
13. thunder
14. heavy
15. loud
16. illumination
17. cost
18. coma
19. *chair*
20. rain
21. reveal
22. moving
23. lion
24. mirror
25. dilatory
26. consider
27. penetrate
28. shake
29. chamber
30. tired
31. blade
32. spell
33. dry
34. careful
35. morsel
36. scalpel
37. death
38. niche
39. rapt
40. glaze
41. roll
42. sink
43. nail
44. rise
45. practice
46. organ
47. pig
48. wind
49. discord
50. dome
51. back
52. listen
53. sure
54. snap
55. twist
56. dark
57. blunt
58. confound
59. stilted
60. operation
61. jump
62. echo
63. wooden
64. prostrate
65. indent
66. glare
67. close
68. immerse
69. *double*
70. express
71. lock
72. boil
73. contain
74. scalp
75. remove
76. route
77. disclose
78. ruin
79. carpet
80. fourth
81. minute
82. sheet
83. hit
84. say
85. local
86. bolt
87. radio
88. olive
89. but
90. sand
91. cup
92. reline
93. pillow
94. whirl
95. the
96. replace
97. hide
98. light
99. snow
100. hole
101. cow
102. breeze
103. other
104. sword
105. sea
106. dog
107. insert
108. throw
109. exit
110. intense
111. stop
112. fade
113. nude
114. haze
115. liver
116. dip
117. triple
118. sound
119. *mention*
120. foot
121. lightning
122. pedal
123. toss
124. rigid
125. confuse
126. roar
127. can
128. daze
129. expose
130. elephant
131. glass
132. linger
133. bounce
134. entrance
135. fresh
136. razor
137. pronounce
138. wet
139. careless
140. bit
141. extreme
142. crevice
143. awe
144. cement
145. raise
146. key
147. fry
148. hair
149. road
150. destroy
151. clay
152. eat
153. fog
154. therefore
155. log
156. way
157. quiver
158. hop
159. forceps
160. stark
161. seek
162. acid
163. tree
164. quick
165. shut
166. water
167. rush
168. have
169. *hospital*
170. certain
171. button
172. mountain
173. sew
174. rubber
175. bottom
176. cat
177. hammer
178. fall
179. sombre
180. restore
181. air
182. saw
183. hail
184. peg
185. roof
186. here
187. bite
188. insane
189. want
190. about
191. warp
192. hear
193. puncture
194. and
195. raving
196. between
197. resound
198. funeral
199. sharp
200. flounder

**Words in List *B* are presented in the order in which they were presented to *S*.

a frequency count which falls within the range from 70 to 100 for a given one of the six concepts and falls within the range from 0 to 30 for the remaining five concepts (with a few exceptions, where it was necessary to go slightly lower than a 70 count to obtain sufficient stimulus material for the experiment). List *B* (also shown in Table 1) contains words (1-78) having frequency counts which fall within the range from 0 to 30 for all six concepts. Words 79 through 200 were obtained from a supplementary scaling study designed specifically to yield words having low frequency counts, and all words fall within the range from 0 to 30. This supplementary scaling study identical in procedure to the initial scaling (3) was necessary, as in the original study only 78 of the 600 words had frequency counts falling within the range from 0 to 30 for all six concepts. In brief, List *A* is comprised of words frequently judged as belonging to some one of the six concepts, while List *B* is comprised of words infrequently judged as belonging to any of the six concepts, i.e., the two lists represent the two ends of a frequency-checking continuum or scale.

2. *Procedure*

Utilizing the words shown in Table 1, nine different experimental word lists were constructed, each list being presented to a separate group of 50 *S*s. Word List 1 contained the 200 words given in List *B*. The remaining 8 word lists each contained the 200 List *B* words immediately preceded by either 6, 18, 30, 60, 90, 120, 150, or 180 List *A* words randomly arranged, but equally divided among the six concepts. That is, with the six List *A* words one word was selected from each of the six concepts, while with the 180 List *A* words 30 words were selected from each of the six concepts.

Since the scale values for List *A* words fall essentially within a range from 70 to 100, a purely random selection of List *A* words to construct the eight experimental word lists would be unsatisfactory. For example, by chance the average scale values for the economic category List *A* words might be 73 for the 18 List *A* words condition, 97 for the 60 List *A* words condition, and 86 for the 180 List *A* words condition. Therefore, the words utilized in constructing the various length List *A*'s were selected to insure that the average scale value for each given concept was relatively the same for the eight experimental lists.

All nine word lists varying in length from 200 to 380 words were presented to *S*s during regular class periods on mimeographed sheets along with instructions to judge each word as to whether they felt it belonged to none, 1, 2, 3, 4, 5, or 6 of the concepts. The *S*s were told to check the words in sequence, to work carefully, but not to spend too much time on any given word. All conditions were administered to each class, conditions being assigned to *S*s at random. A total of 14 undergraduate psychology classes at New York University was employed and total testing was accomplished during a three-week period. None of the 450 *S*s of the present study had been tested in any of the previous studies cited.

C. RESULTS

Table 2 shows the concept spans, or mean number of words checked as belonging to each of the six conceptual categories, on List *B*, for the nine anchoring conditions. The results indicate that for each of the six conceptual categories, as the number of anchoring stimuli are increased, there is a marked decrement in List *B* checking behavior. For example, for the theoretical concept, a mean number of 32.5 words were checked as theoretical under the zero anchor condition, 22.3 words were checked under the 10-anchor condition, and 15.9 words were checked under the 30-anchor condition. A simple analysis of variance was performed for each conceptual category, and yielded probability values of less than .001 for the economic, aesthetic, political, and religious concepts, while the values for the theoretical and social concepts were approximately .10 and .20, respectively.

Figure 1 presents the average number of words checked on List *B* for all six concepts combined as a function of experimental groups, i.e., the number of anchoring stimuli per concept given in List *A*. The plot shows a marked drop in List *B* checking behavior after exposure on List *A* to only one anchoring stimulus per conceptual category. Also, it would appear that 10 anchoring stimuli per category is sufficient to produce the maximum decrement in List *B* checking behavior, as after this point, no further drop is seen, at least, in the range investigated in this study.

Table 3 gives the concept spans for List *A*, under the various anchoring conditions. Since List *A* is composed exclusively of words showing previously high scaled values of inclusion in the categories, it was expected that the mean number of words checked per category should approximately equal the number of words or anchoring stimuli present. In general, the results conform to this expectation with, for example, the political category revealing a high degree of correspondence between the number of anchoring stimuli given (i.e., 1, 3, 5, 10, 15, 20, 25, and 30) and the mean number checked (i.e., 1.0, 3.2, 5.4, 12.4, 14.1, 22.6, 27.7, and 32.9).

Having analyzed the general decremental effect of varying numbers of anchoring stimuli on low scaled words, an analysis was made to determine

TABLE 2

MEANS AND *SD*'S OF THE CONCEPT SPANS FOR LIST *B* FOR THE ANCHORING CONDITIONS FOR THE SIX CONCEPTUAL CATEGORIES

Concept anchor	Th.		Ec.		As.		So.		Po.		Re.	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
0	32.5	22.8	43.5	21.9	32.9	20.8	59.3	26.7	36.8	17.3	27.2	15.5
1	26.2	17.8	34.9	19.6	29.3	21.6	59.2	32.7	22.8	16.3	21.4	14.5
3	26.8	23.9	32.7	20.1	24.8	17.5	59.6	32.8	25.5	15.4	21.2	14.7
5	25.0	21.7	35.5	23.8	23.6	18.0	50.5	30.2	20.9	17.6	14.9	11.7
10	22.3	28.7	26.8	27.3	17.3	18.0	46.7	35.8	18.2	17.7	12.3	12.0
15	29.7	35.0	23.5	18.8	21.0	15.3	46.8	31.1	20.7	17.8	15.0	10.0
20	22.3	27.0	19.0	16.2	17.7	19.2	51.0	41.4	17.0	12.5	14.0	13.4
25	28.0	35.0	27.0	26.7	18.9	14.5	56.7	39.1	17.2	14.8	11.7	10.1
30	15.9	19.6	25.5	27.3	14.8	13.4	51.0	35.2	18.0	17.6	14.4	11.5

TABLE 3

MEANS AND *SD*'S OF THE CONCEPT SPANS FOR LIST *A* FOR THE ANCHORING CONDITIONS FOR THE SIX CONCEPTUAL CATEGORIES

Concept anchor	Th.		Ec.		As.		So.		Po.		Re.	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
1	1.1	.8	1.3	.5	1.2	.7	1.3	.9	1.0	.6	1.2	.5
3	3.1	2.2	3.5	1.0	2.8	2.0	3.9	1.6	3.2	.9	3.3	1.1
5	4.9	2.4	6.1	1.9	3.8	1.9	7.4	2.2	5.4	1.6	5.5	1.5
10	9.5	4.3	12.9	3.8	8.8	4.5	16.2	5.1	12.4	1.5	11.7	3.7
15	13.0	3.7	16.9	3.9	13.1	4.4	18.7	18.8	14.1	3.5	15.8	2.8
20	19.7	6.1	24.5	3.8	18.4	7.2	30.1	10.8	22.6	4.8	22.5	6.0
25	24.9	12.3	32.0	6.4	23.8	9.3	35.7	13.5	27.7	3.3	28.2	6.9
30	25.7	10.8	39.7	13.8	24.0	12.8	47.7	24.7	32.9	2.8	36.7	13.6

how this checking decrement was distributed throughout the range of the 200 low scaled List *B* words. The results are shown in two graphs. Figure 2 is a plot of the average number of words checked for all six concepts combined on List *B* as a function of successive word groups for the various anchoring conditions. Specifically, List *B* was divided into five successive 40-word groups and the average number of words checked for all six concepts combined was determined for each anchoring condition for the successive word groups. It is apparent from Figure 2 the checking decrements produced by the various anchoring conditions are relatively uniform throughout the range of List *B*, as evidenced by the general paralleling of the curves. The corresponding plots for each separate conceptual category are of the same form as the composite in Figure 1.

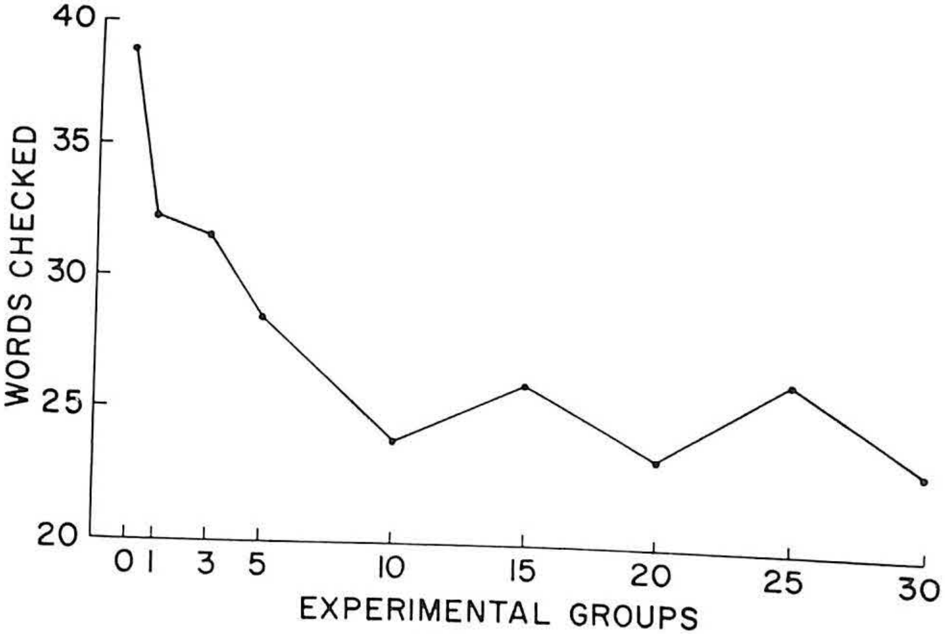


FIGURE 1
A PLOT OF THE AVERAGE NUMBER OF WORDS CHECKED FOR ALL SIX CONCEPTS COMBINED ON LIST *B* AS A FUNCTION OF THE NUMBER OF ANCHORING STIMULI

Figure 3 represents a plot of the average percentage of *S*s checking individual words 1-10 and 191-200 for all six concepts combined on List *B* for three particular anchoring conditions. These three particular anchoring conditions were chosen to tap the two extremes of the anchoring variable and that single value having the relatively greatest proportional effect, namely one anchor per category. The most prominent feature seen is the general paralleling of the three curves, with the decremental effect acting uniformly on individual words. To conserve space, the remaining words of List *B* are not shown, nor the plots for each separate category, however, the results in general correspond in form to those shown.

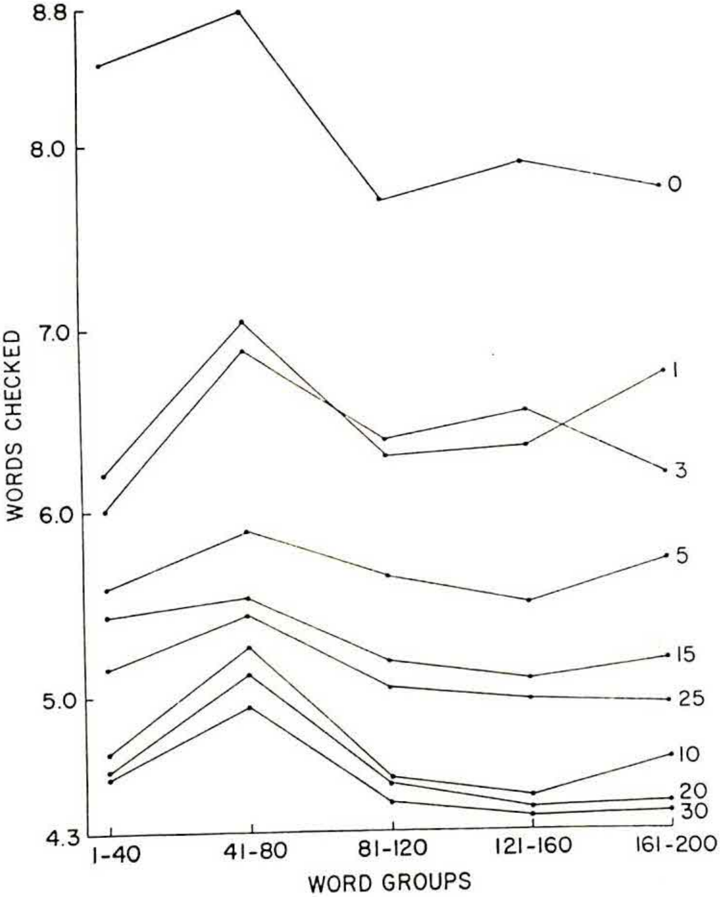


FIGURE 2

A Plot of the Average Number of Words Checked for All Six Concepts Combined On List *B* As a Function of Successive Word Groups for the Nine Anchoring Conditions

D. Discussion

The present investigation has attempted to examine the effects of varying numbers of high scaled anchoring stimuli on the judgment of subsequent low scaled stimuli. Its starting point was a previous study (2) in which a judgmental model for treating verbal conceptual or categorizing behavior was developed. Such a model, while predicting in general that increasing the number of anchoring stimuli should produce the decremental effects found, could not specify in detail the actual shape of the function, as shown in Figure 1. However, data from an earlier study (4) is relevant to this issue. Briefly, *S*s were asked to judge one series of lifted weights and then an expanded series. Varying amounts of practice trials (1, 4, 8, 12, and 28) were given for the first series before shifting to the second. The results showed that the greater the amount of practice with the anchoring stimuli, the more slowly does the scale of judgment shift to its new position, with some indication that the smallest practice condition (one trial) had the greatest proportional effect. This latter finding is one of the more significant aspects of the present study, namely, that only one anchoring stimuli per conceptual category produced the greatest proportional decremental effect. Further, the finding of the present study that the decremental effect is apparently maximized at 10 anchoring stimuli per category, also, is in agreement with the earlier study (4), in that here the effective shifting of the scale seemed to maximize and stabilize around 8 to 12 practice trials. The apparent agreement between these two studies, however, must be interpreted with caution, because of different procedures, analysis techniques, and stimuli being judged. Additional work is needed to confirm that the shape of the function produced by anchoring stimuli is the same whether the stimuli are weights or words.

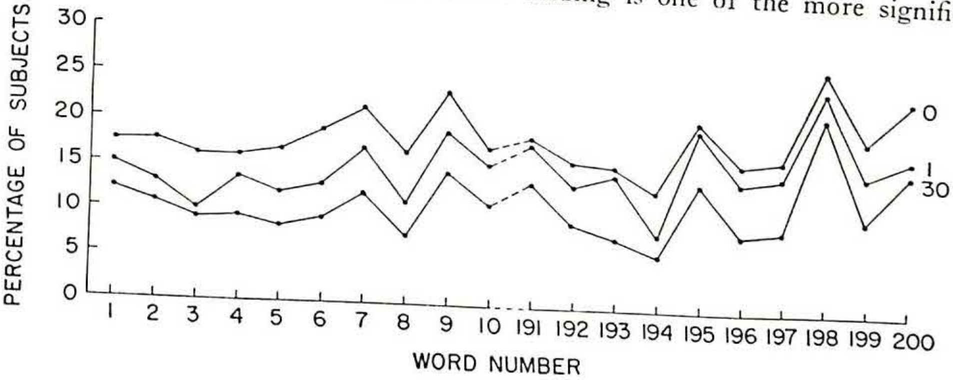


FIGURE 3

A PLOT OF THE AVERAGE PERCENTAGE OF *S*s CHECKING INDIVIDUAL WORDS 1-10 AND 191-200 FOR ALL SIX CONCEPTS COMBINED ON LIST *B* FOR A SELECTED THREE ANCHORING CONDITIONS

Having noted some points of possible agreement between the weight study and the present word study, one area in which the two studies radically diverge should be examined. Figures 2 and 3 show in a pronounced way the extreme stability and consistency of the various anchoring conditions, manifesting their effect in a uniform and systematic pattern throughout the entire range of the 200 List *B* words. In contrast, in the weight study, the effects of varying amounts of practice or exposure to the anchoring stimuli were quickly reduced. Specifically, the effect was completely dissipated by the sixth stimulus presentation, whereas with the words the effect was as strong on the last of the 200 stimulus words as on the first.

This difference in persistency of anchoring effects, between weight stimuli and word stimuli, might reside in the more stabilized judgmental scale structure associated with verbal conceptual categorizing developed by the extensive prior learning experiences inherent in language acquisition. Two questions now arise, (*a*) "Just how persistent and stable is the anchoring effect produced by word stimuli?" and (*b*) "What learning processes and variables are associated with the establishment of a stabilized judgmental scale structure relevant to verbal conceptual categorizing behavior?"

E. SUMMARY

The present study was directed toward discovering the optimum number of discrete stimuli needed to produce the greatest anchoring effect for a series of word-concept scales. Specifically, words previously scaled high as to their frequency of inclusion in given conceptual categories served as anchoring stimuli, followed by a series of low scaled words. The number of anchoring stimuli was systematically varied from zero to 30 per conceptual category, with the results indicating that one anchor per category produces proportionately the greatest effect (i.e., decrements in checking the low scaled words) and that 10 anchors per category maximizes the effect, over the range investigated. Further it was shown that this anchoring effect acts uniformly and systematically throughout the entire range of low scaled words, producing uniform checking decrements on individual words.

REFERENCES

1. MAYZNER, M. S., & TRESSELT, M. E. Concept span as a composite function of personal values, anxiety, and rigidity. *J. Pers.*, 1955, **24**, 20-33.
2. ———. Shifts in connotative meaning of words as a function of previous restrictive experience. *J. Exp. Psychol.*, 1958, **55**, 200-205.
3. PRITCHARD, L., MAYZNER, M. S., & TRESSELT, M. E. The Mayzner-Tresselt word-concept check list. *Psychol. Newsl.*, 1955, **7**, 27-33.
4. TRESSELT, M. E. The influence of amount of practice upon the formation of a scale of judgment. *J. Exp. Psychol.*, 1947, **37**, 251-260.
5. TRESSELT, M. E., & MAYZNER, M. S. Consistency of judgments in categorizing verbal material. *Psychol. Rep.*, 1958, **4**, 415-421.

15 Sydney Avenue
Deal, New Jersey

Department of Psychology
New York University
New York 3, New York

Published as a separate and in *The Journal of Psychology*, 1959, **47**, 117-125.

ANAGRAM SOLUTION TIMES: A FUNCTION OF TRANSITION PROBABILITIES*

Bell Telephone Laboratories and New York University

M. S. Mayzner and M. E. Tresselt

A. Problem

It may be assumed that anagram solution times are a function of many factors. A previous study (13) demonstrated the effects of two variables, namely, anagram letter order (i.e., the rearrangement of the letters from their correct position in the word) and word frequency, on anagram solution times. Specifically, the "similarity" of anagram letter order to word letter order was either maximized (easy order) or minimized (hard order) and four levels of Thorndike-Lorge word frequency counts from frequent to infrequent were chosen for investigation. The results showed that hard letter orders and low word frequency counts both produced significantly marked increases in anagram solution times, in contrast to easy letter orders and high word frequency counts.

The present investigation examines the effect of different transition probabilities on anagram solution times. In any given anagram or letter array there is a transition probability or digram frequency associated with each sequential pair of letters, and digram frequency tables are available (15) which permit the relative assignment of a numerical value to this frequency. If the digram frequencies for all sequential pairs of letters of a given anagram are summed, one obtains a transition probability total for the anagram, which can be made to vary by permuting the letters of the anagram. A given letter set then can be arranged to form an anagram with a high transition total or a low transition total. Assuming now that anagram solving involves permuting the letters of the anagram and further that letter pairs with low digram frequencies are easier to permute than those with high digram frequencies, then in general, anagrams with low transition totals should be solved faster than those with high transition totals. This hypothesis is tested in the present study.

B. Method

1. *Subjects*

Forty *Ss* from introductory psychology classes at New York University were employed in the experiment, 20 in each of the two major conditions.

---

*Received in the Editorial Office on November 6, 1958, and published immediately in Provincetown, Massachusetts. Copyright by The Journal Press.

2. *Material*

The stimulus material consisted of 20 five-letter anagrams typed in caps on plain white 3- x 5-in. index cards. The 20 words from which the anagrams were constructed were selected from the Thorndike-Lorge word list (19), to provide four groups of five words each, differing in frequency of occurrence as follows: Group *VF* (very frequent) was characterized by words occurring 100 times or over per million, Group *F* (frequent) by words occurring at least 50 times per million and not so many as 100 times per million, Group *I* (infrequent) by words occurring at least once per million and not so many as twice per million, and Group *VI* (very infrequent) by words occurring less than once per million but more than once per four million. The words, numbered in order of presentation to *S*s, are presented in Table 1, along with their two anagram forms (i.e., high transition totals and low transition totals with the transition totals in parenthesis following the anagrams). The 20 stimulus words from which the anagrams were constructed are the same as those used in the previous anagram study (13), to

TABLE 1
WORDS AND ANAGRAMS EMPLOYED IN THE EXPERIMENT

Words	Anagrams Low transition totals		High transition totals	
	Group VF			
1. chair	ihrca	(28)	hicar	(127)
2. sugar	gsrua	(16)	rasgu	(71)
3. train	irtna	(49)	antir	(190)
4. party	rpyta	(31)	atryp	(108)
5. labor	rbloa	(16)	arolb	(86)
	Group F			
6. beach	ehcba	(12)	hecab	(177)
7. model	dmloe	(18)	lomed	(116)
8. clerk	ekrlc	(4)	reckl	(118)
9. fault	tflua	(14)	latuf	(100)
10. giant	igtna	(23)	tinga	(163)
	Group I			
11. patio	otpia	(33)	aitop	(118)
12. cobra	aocrb	(13)	baroc	(86)
13. roach	aohrc	(10)	hocar	(80)
14. baton	obtna	(17)	tanbo	(131)
15. jaunt	ajtnu	(1)	jutan	(147)
	Group VI			
16. tango	gtnao	(15)	gaton	(185)
17. groin	rnogi	(24)	rigon	(122)
18. peony	oeypn	(10)	enopy	(85)
19. ghoul	lhguo	(2)	hugol	(35)
20. triad	rtdai	(44)	datir	(137)

provide continuity and to make possible direct comparisons between anagram solution times under the different conditions investigated in the two studies.

3. *Experimental Design*

The *Ss* were assigned randomly to the two major conditions which involved presenting the same 20 words as either anagrams having low transition totals or as anagrams having high transition totals (as shown in Table 1). The transition total was obtained by determining for each word, the digram frequencies for all 16 distinct letter pairs possible with a five-letter word, excluding the four-letter pairs of the word itself, and then so arranging the letters that the four sequential digram frequencies summed for these letter pairs, would either minimize (low transition total) or maximize (high transition total). For example, with chair (*ihrca*) the four digram frequencies for the letter pairs *ih, hr, rc,* and *ca* are 0, 3, 5, and 20 respectively, summing to 28, while with chair (*hicar*) the four digram frequencies for the letter pairs *hi, ic, ca,* and *ar* are 43, 24, 20, and 40, respectively, summing to 127. Since letter order was shown to be a significant variable influencing anagram solution times, a further constraint was introduced in constructing the anagrams. To control for this variable, all anagrams both in the low and high transition conditions involved hard letter order arrangements.

4. *Procedure*

Subjects were told that a five-letter anagram is simply five scrambled letters which when properly rearranged will make a single five-letter word. After being shown an example, they were informed that they would have 20 anagrams to solve, and if they could not find a solution in four minutes, the correct solution would be given to them, before going on to the next one. Solutions were given by *S* verbally (i.e., *S* had no paper or pencil) and solution times were recorded with a stop watch. All *Ss* were tested individually.

C. RESULTS

Available for purposes of analysis were 20 time scores for each *S*, with each score having a possible range from 1 to 240 seconds. Because of the nature of the distributions of the five time scores for each word frequency group (*VF, F, I,* and *VI*)—transition probability total (low and high) combination for each *S*, the *S*'s median time score for each combination was selected as most representative of his performance. Thus for each *S*, four time scores were made available and used in analyzing the effect of the two major variables. Further, since the distribution of these median time scores revealed marked heterogeneity of variance between the various experimental treatments, non-parametric techniques were used throughout.

Table 2 gives the major results of the study, showing means (*M*), standard deviations (*SD*), medians (*MD*), and ranges (*R*) for the time scores. Both means and medians reveal a steady increase in anagram solution times

TABLE 2

ANAGRAM SOLUTION TIMES (IN SEC.) FOR TRANSITION PROBABILITY TOTAL AND WORD FREQUENCY CONDITIONS (BASED ON MEDIANS)

	M	*SD*	*MD*	*R*
Low transition total				
Group *VF*	70.8	78.7	54.0	2-240
Group *F*	50.1	63.0	28.5	5-230
Group *I*	92.5	66.7	65.5	14-240
Group *VI*	97.1	81.9	69.0	5-240
High Transition total				
Group *VF*	121.8	98.8	76.5	4-240
Group *F*	101.0	79.1	77.5	20-240
Group *I*	183.4	71.6	229.0	54-240
Group *VI*	198.6	78.1	240.0	24-240

Note—Time scores from which *M, SD, MD,* and *R* are calculated are the median solution times for each frequency group of five anagrams.

going from the low transition total—*VF* condition to the high transition total—*VI* condition, except for slight reversals between the *VF* and *F* word groups. Within this design several comparisons are possible. To evaluate the effect of the transition probability variable, comparisons of the low versus high transition total conditions for each of the four word frequency groups, were carried out employing the Mann-Whitney *U* Test (11). These four comparisons for the *VF, F, I,* and *VI* word groups yielded probability values of .10, less than .01, less than .001, and less than .001, respectively, thus indicating a significant effect due to this variable, in line with the stated hypothesis.

In evaluating the effect of the word frequency variable, two separate analyses were made. Under each transition probability total condition (i.e., low, high), the Friedman Two-Way Analysis of Variance by Ranks test (2) was employed. The probability values associated with both word frequency group comparisons were well below the .001 level, in agreement with the results of the previous study cited (13).

D. DISCUSSION

The results of the present study indicate the significant effect of transition probabilities on anagram solution times. They demonstrate how this concept, first fully dealt with in this context by Shannon (18), can be used to quantitatively vary the statistical structure of language and to observe the effects of such variation on subsequent behavior. Relevant to this issue are several anagram studies by Hollingworth (4), Devnich (1), and Nissenson and Sargent (14, 17), who on the basis of Gestalt laws of organization hypothesized that well patterned letter groups (e.g., *intellectual* presented as *"all in lettuce"*) should be more difficult to reorganize than loosely patterned letter groups (i.e., *intellectual* presented as *"alinletluetc"*). The results of these several experiments were contradictory. Devnich obtained positive results in favor of the hypothesis, but the other studies obtained mainly negative results. Johnson has concluded therefore that "the weight of the evidence at present suggests that there is no difference between well-patterned and loosely patterned letter groups in respect to ease of reorganization, but Devnich's experiment cannot be disregarded" (5, pp. 230-231). The findings of this experiment suggest that well patterned letter groups are, in fact, more difficult to reorganize than loosely patterned groups, if patterning is defined in terms of transition probabilities.

It was stated in the previous study (13) that an anagram constitutes a problem whose stimulus properties tend to evoke a variety of implicit responses, and the manner in which letter order and word frequency variables affect this production process was discussed. It is now suggested that transition probability or digram frequencies may also affect this production process in the following fashion. Some of the implicit hypotheses generated are two-letter groupings having high digram frequencies, which are presumably well learned habit structures inherent in the statistical structure of the language, and which constitute "fine" structures of the word itself. Further, the generation of such digram "fine" structures proceeds more rapidly with a letter array in which there is little transition bind between the sequential letters (the low transition total condition) allowing for easy recombination than with a letter array in which there is already considerable high transition bind between the sequential letters (the high transition total condition) hindering easy recombination. Thus, with the more rapid production of digram "fine" structures in the low transition total condition, there is presumably a greater chance of matching that letter array which is, in fact, the word-solution.

Since the 20 words employed in this study were also utilized in the previous study (13), it is possible to make direct comparisons between the studies. Table 3 presents the median solution times for each individual word in the two studies. The transition totals are given also for the four anagram forms of the words used. A direct comparison can be made in the table between the solution times for each word in the low and high transition total

TABLE 3

ANAGRAM SOLUTION TIMES (IN SEC.) AND TRANSITION TOTALS FOR EACH INDIVIDUAL WORD UNDER EASY AND HARD LETTER ORDER AND LOW AND HIGH TRANSITION PROBABILITY TOTAL CONDITIONS

Words	Transition total	Easy letter order	Hard letter order	Transition total	Transition total	Low transition total	High transition total	Transition total
Group *VF*								
1. chair	(131)	10.0	12.5	(38)	(28)	8.5	57.0	(127)
2. sugar	(72)	10.5	91.5	(29)	(16)	182.5	88.5	(71)
3. train	(108)	5.0	8.0	(103)	(49)	12.0	118.5	(190)
4. party	(42)	14.0	24.0	(46)	(31)	51.0	65.0	(180)
5. labor	(78)	5.0	210.0	(46)	(16)	159.5	201.0	(86)
Group *F*								
6. beach	(54)	3.0	9.5	(61)	(12)	8.5	37.5	(177)
7. model	(57)	4.5	191.5	(55)	(18)	62.5	59.0	(116)
8. clerk	(136)	17.5	22.0	(33)	(4)	31.5	91.5	(118)
9. fault	(25)	7.0	36.0	(93)	(14)	56.0	136.5	(100)
10. giant	(58)	7.5	154.0	(59)	(23)	29.0	240.0	(163)
Group *I*								
11. patio	(127)	22.0	125.0	(34)	(33)	71.0	218.0	(118)
12. cobra	(53)	50.5	240.0	(16)	(13)	240.0	240.0	(86)
13. roach	(54)	9.5	19.5	(52)	(10)	38.0	80.5	(80)
14. baton	(122)	9.5	194.0	(104)	(17)	85.0	240.0	(131)
15. jaunt	(47)	5.0	37.5	(83)	(1)	33.0	240.0	(147)
Group *VI*								
16. tango	(132)	45.0	240.0	(77)	(15)	240.0	240.0	(185)
17. groin	(148)	31.0	35.0	(24)	(24)	25.5	240.0	(122)
18. peony	(88)	12.0	240.0	(72)	(10)	45.5	240.0	(85)
19. ghoul	(62)	19.5	33.5	(4)	(2)	75.5	240.0	(35)
20. triad	(86)	5.0	103.0	(95)	(44)	61.5	240.0	(137)

Note—Transition probability totals for the anagram form of the word used appear in parenthesis before and after the anagram solution times obtained for the different experimental conditions (i.e., the easy and hard letter order conditions of the previous experiment cited and the low and high transition probability total conditions of the present study).

conditions of this study and reveal that the solution times for 16 words are in favor of the hypothesis, two are against (*sugar* and *model*) and two are tied (*cobra* and *tango*). In the previous study, no initial count of transition totals was made, but a count made after the present study was carried out, on the hard letter order anagrams shows that the counts for most of the words fall approximately midway between the low and high totals of the present study. This fortuitous outcome now makes it possible to make additional comparisons between anagram solution times as related to transition totals. A comparison of the low transition total column solution times with the hard letter order column solution times, reveals that the solution times for 11 words are in favor of the hypothesis, seven are against, and two are tied, while a comparison of the high transition total column solution times with the hard letter order column solution times, reveals that the solution times for 14 words are in favor of the hypothesis, three are against, and three are tied. The results taken together suggest that the effect of the transition probability variable is sensitive enough to produce differences in anagram solution times for low, medium, and high transition totals, but the effect is strongest when the difference between transition totals is maximized, as in the present study. The apparent lack of effect of the transition probability variable on solution times when easy letter orders are employed is also shown in Table 3. This latter finding coupled with the rather extreme range of solution times found in Table 3 for any given word frequency, letter order, or transition probability combination implies that the three variables so far studied far from exhaust the variance associated with anagram solving.

In addition to the issues already discussed, certain methodological considerations evolving from these two studies warrant brief comment. For example, with the identification of letter order, transition probability, and word frequency variables as major sources of variation in anagram solution times, it becomes possible to construct a wide range of anagrams differing in difficulty level, from very easy to very hard. The construction of such an anagram list scaled for difficulty level, could function in problem solving studies, in an analogous fashion, to the various lists of nonsense syllables, scaled for association value (3) in learning studies. Further, many studies (6, 7, 8, 9, 10, 12, 16, 17) employing anagrams as problem material have universally neglected to control for letter order, transition probabilities, or word frequency variables in any systematic fashion. In general, the results of these studies may be confounded by the effects of these uncontrolled variables. In particular, the many studies which have tested set phenomena through the use of anagrams possessing double and multiple solutions, reveal that the majority of these double and multiple solutions are words varying markedly in their Thorndike-Lorge frequency counts. Quite probably, such anagrams do not possess equal response potentials for all solutions and could confound the effect of set factors with unbalanced word frequency factors. Letter order and transition probability variables could also play an analogous rôle to word frequency variables in mutiple solution anagrams. Controlling for all these variables in multiple solution anagrams would probably severely limit the number of words that could be found possessing these characteristics, but with these variables controlled unsuspected relationships might emerge.

E. SUMMARY

A study was made of the effect of transition probability and word frequency variables on anagram solution times. Two levels of transition probability and four levels of word frequency were examined. Transition probability was varied by either maximizing or minimizing digram frequency totals to yield for each word employed an anagram having a low transition probability total and an anagram having a high transition probability total. Word frequency was varied by selecting words varying in their Thorndike-Lorge frequency counts from very frequent to very infrequent.

The results clearly showed the highly significant effect of these two variables on anagram solution times. Specifically, high transition probability totals and low word frequency counts both produced significantly marked increases in anagram solution times, in contrast to low transition probability totals and high word frequency counts.

REFERENCES

1. DEVNICH, G. E. Words as "Gestalten." *J. Exp. Psychol.*, 1937, **20**, 297-300.
2. FRIEDMAN, M. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *J. Amer. Statis. Ass.*, 1937, **32**, 675-701.
3. GLAZE, J. A. The association value of nonsense syllables. *J. Genet. Psychol.*, 1928, **35**, 255-267.
4. HOLLINGWORTH, H. L. Verbal Gestalt experiments with children. *J. Exp. Psychol.*, 1938, **23**, 90-95.
5. JOHNSON, D. M. The Psychology of Thought and Judgment. New York: Harper, 1955.
6. MALTZMAN, I., & MORRISETT, L., JR. Different strengths of set in the solution of anagrams. *J. Exp. Psychol.*, 1952, **44**, 242-246.
7. ———. The effects of single and compound classes of anagrams on set solutions. *J. Exp. Psychol.*, 1953, **45**, 345-350.
8. ———. Effects of task instructions on solution of different classes of anagrams. *J. Exp. Psychol.*, 1953, **45**, 351-354.
9. MALTZMAN, I., FOX, J., MORRISETT, L., JR. Some effects of manifest anxiety on mental set. *J. Exp. Psychol.*, 1953, **46**, 50-54.

10. Maltzman, I., Eisman, E., Brooks, L. O., & Smith, W. M. Task instructions for anagrams following different task instructions and training. *J. Exp. Psychol.*, 1956, **51**, 418-420.

11. Mann, H. B., & Whitney, D. R. On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Statis.*, 1947, **18**, 50-60.

12. Mayzner, M. S., & Kendler, H. H. The influence of rigid and flexible training methods. *Amer. Psychol.*, 1954, **9**, 428.

13. Mayzner, M. S., & Tresselt, M. E. Anagram solution times: A function of letter order and word frequency. *J. Exp. Psychol.*, 1958, **56**, 376-379.

14. Nissenson, M., & Sargent, S. S. Words as configurations. *J. Exp. Psychol.*, 1941, **28**, 85-89.

15. Pratt, F. Secret and Urgent. New York: Blue Ribbon Books, 1942.

16. Rees, H. J., & Israel, H. E. An investigation of the establishment and operation of mental sets. *Psychol. Monog.*, 1935, **46**, No. 6 (Whole No. 210).

17. Sargent, S. S. Thinking processes at various levels of difficulty. *Arch. of Psychol.*, 1940, No. 249.

18. Shannon, C. E. A mathematical theory of communication. *Bell Syst. Tech., J.*, 1948, **27**, 379-423, 623-656.

19. Thorndike, E. L., & Lorge, I. The Teacher's Word Book of 30,000 words. New York: Bureau of Publications, Columbia Univ., 1944.

15 Sydney Avenue
Deal, New Jersey

Department of Psychology
New York University
New York 3, New York

Published as a separate and in *The Journal of Psychology*, 1959, **47**, 127-136.

SEX DIFFERENCES IN INFERRING PERSONALITY TRAITS*

Harvard University

David Shapiro and Renato Tagiuri[1]

A. Purpose

In our daily social interactions much of what we feel and do is based on our perceptions and interpretations of what we observe in the other person. Yet, we never have "direct" evidence on each facet of the other individual. Rather, we rely a great deal on our capacity to extrapolate from available "information" to other unknown aspects of the person.

The basis of our impression may vary a great deal in nature and usefulness. Knowledge about a person's *rôle*, for example, is very helpful for making a good many predictions of a certain type. Sometimes the information we have about a person is in the form of one or more traits. We may be told that a certain individual is *generous,* and from this we guess what other traits he might have. We may, for instance, feel quite certain that he will also be *thoughtful.* On the other hand, we might find it difficult to predict his *honesty* from his *generosity*. By the process of inference, then, we fill in what we need to know in order to act, and our inferences are guided by a more or less implicit "theory" of personality.

No doubt there is a great deal of agreement within any cultural group as to what inferences can be made about a person from various forms of limited information. Common experiences foster certain agreed-upon "theories" of personality. Yet, clearly, individuals often disagree, given certain information, as to their conclusions about a person. For many, *intelligent* people are likely to be *witty* and *warm*; for others, the same trait leads to different conclusions—*intelligent* means *sarcastic* and *cold.*

In this paper, we inquire as to whether the different personal experiences of men and women affect their social perceptions. More specifically, we ask whether traits have somewaht different meanings for men and women.

---

*Received in the Editorial Office on November 15, 1958, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This work was done in the Laboratory of Social Relations and Graduate School of Business Administration, under Office of Naval Research Contract No. N5 ori 07670, and also in the Department of Psychiatry, Harvard Medical School, under NIMH Training Grant No. 2M-6378 C2.

B. PROCEDURE

The meaning of a trait was defined operationally as the set of inferences made from a *given* trait to a set of other traits (1). Concretely, a subject was given a statement such as "People who are INTELLIGENT" and was asked to indicate whether such people:

very often are,
tend to be,
may or may not be,
tend not to be,
seldom are,

aggressive, active, etc., for a list of 59 such traits (see Table 1).[2] Four different traits, INTELLIGENT, CONSIDERATE, INDEPENDENT, and INCONSIDERATE were used, one at a time, as the *given* trait.

Sixty male and 60 female undergraduate liberal arts students from the Boston area were tested with each of the four *given* traits—a total of 480 subjects. *One and only one* of the four *given* traits was used with a subject.

C. RESULTS

1. *Inferences from Given Traits to List Traits*

Let us compare men and women in terms of the inferences they make from the *given* traits to the *list* traits (Table 1). Men and women show remarkable agreement in their distribution of responses to each *given* trait. Chi-square tests comparing men and women for their distribution of responses on each *list* trait, yield results within chance expectancy. There is also agreement as to the dominant sign of the inference in almost all cases. A few exceptions are found, as would be expected, in cases where the majority of responses are "may or may not be," or where there is *within* sex disagreement as to the direction of inference. In essence, however, the meaning of each of the four traits used in this study is quite comparable for men and women.

Notice that in the rating scale used, two steps in inference strength or certainty are distinguishable for positive inferences and two for negative ones. The inference "may or may not be" represents an indefinite or uncertain connection between *given* and *list* trait. In a sense, for a subject, a particular response represents the degree of strength or certainty or subjective likelihood that, given one trait in a person, a specific other trait will be present or absent.

---

[2] Two alternate forms of the trait list were used; one listed the traits in roughly alphabetical order, the second, in the opposite order. Inasmuch as differences due to form were minor, data from both forms were pooled for purposes of this paper.

Since there is almost complete agreement between the two sexes as to the direction of the association between *given* and *list* traits, we shall next distinguish simply between *definite* inferences ("very often are," "tend to be," "tend not to be," and "seldom are" responses) and *indefinite* inferences ("may or may not be" responses). We can now compare men and women for the average *number* of definite inferences they make, from a specific *given* trait, to all the traits in the list (Table 2).

Men and women consistently respond very similarly, again. Yet, women make more definite inferences than men, for each of the *given* traits, although the differences are slight and statistically not significant even when the data are combined. The largest difference occurs for inferences made from the *given* trait CONSIDERATE, where women, on the average, make 3.3 definite inferences more than men, out of a total of 53 possible.[3] Thus, women are slightly more willing, from knowledge of a *given* trait, to commit themselves as to the *presence* or *absence* of a *list* trait. However, individual differences are great and far outweigh group differences. (The average standard deviation of the measure of inference readiness is 13.1; the average group difference is 2.5.)

If only the more *extreme* definite inferences are considered ("very often are" and "seldom are"), again disregarding the signs, the findings parallel those already reported for all definite inferences (Table 3). However, the difference attains significance in the case of the *given* trait CONSIDERATE where women, on the average, make 3.3 more extreme inferences than men. In addition, when the probabilities of the differences for the four *given* traits are combined they yield a value of less than .01. Comparison of the data in Tables 2 and 3 shows that the slight but consistent sex differences are accounted for entirely by the *extreme* definite inferences. Thus, women differ from men only in that they consistently give a greater number of *extreme* definite inferences. Inferences "tend to be" and "tend not to be" are equally frequent for the two sexes.

Again we need point out that individual variability is great (average standard deviation is 8.0), giving far less weight to group differences (average group difference is 2.5) and more to personal predilections in the readiness to make extreme trait inferences.

---

[3] Inferences to six of the 59 traits listed—*intelligent, unintelligent, considerate, inconsiderate, independent,* and *dependent* were omitted in these counts since they were used also as the *given* traits.

TABLE 1
Inferences Made from Each Given Trait to Each List Trait*

		Intelligent					Considerate					Independent					Inconsiderate				
		1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Aggressive	M	4	15	36	4	1	1	0	32	22	5	12	23	22	3	0	13	21	26	0	0
	F	5	11	38	4	2	2	3	18	32	5	16	24	17	3	0	18	24	17	0	1
Active	M	14	27	19	0	0	7	14	38	1	0	12	27	18	2	1	3	16	35	6	0
	F	10	24	25	1	0	6	15	37	2	0	24	23	12	0	1	3	9	40	8	0
Awkward	M	1	3	32	22	2	1	1	41	11	6	0	0	35	19	6	3	8	41	8	0
	F	0	2	27	24	7	1	1	39	12	7	0	1	28	23	8	3	12	37	8	0
Apathetic	M	1	2	10	30	17	1	1	16	21	21	1	3	22	20	14	5	8	34	12	1
	F	1	3	10	30	16	1	3	16	22	18	1	4	13	20	22	9	12	25	11	3
Boastful	M	0	8	41	6	5	0	0	13	30	17	5	18	27	7	3	6	30	22	2	0
	F	1	8	36	13	2	1	0	10	34	15	6	11	36	4	3	11	19	25	5	0
Clever	M	16	31	13	0	0	1	11	46	1	1	5	26	28	0	1	1	9	41	6	3
	F	20	29	10	1	0	3	8	46	3	0	8	26	26	0	0	3	6	45	4	2
Considerate	M	3	20	31	5	1	50	8	1	0	1	1	13	30	15	1	0	2	3	8	47
	F	7	16	37	0	0	56	3	1	0	0	3	10	38	9	0	1	0	2	15	42
Cold	M	1	8	40	8	3	0	0	6	29	25	2	13	37	6	2	9	27	19	5	0
	F	1	10	38	8	3	0	0	7	30	23	2	16	32	9	1	13	29	14	3	1
Courteous	M	7	20	31	2	0	28	27	5	0	0	2	13	38	7	0	2	0	8	33	17
	F	9	23	28	0	0	41	17	2	0	0	2	9	45	4	0	1	1	15	30	13
Unreliable	M	3	1	27	20	9	1	1	21	16	21	0	5	31	15	9	5	23	24	6	2
	F	0	1	13	35	11	0	3	8	24	25	1	4	25	18	12	15	20	15	7	3
Unsympathetic	M	1	6	25	25	3	1	2	2	22	33	1	15	31	9	4	13	28	11	5	3
	F	0	3	30	21	6	2	2	1	17	38	1	11	38	8	2	20	24	6	8	2
Cheerful	M	0	9	46	5	0	7	32	19	1	1	2	13	42	3	0	0	1	35	22	2
	F	4	8	47	0	1	19	28	12	1	0	2	20	36	2	0	2	4	37	15	2
Deliberate	M	13	27	16	4	0	4	15	38	3	0	4	34	14	7	1	2	13	29	12	4
	F	10	24	19	6	1	3	18	32	4	3	17	31	8	2	2	9	10	26	13	2
Dishonest	M	1	0	33	17	9	0	1	18	29	12	1	1	38	12	8	0	9	48	3	0
	F	0	0	26	23	11	0	1	12	23	24	0	0	38	11	11	1	4	49	2	4
Dependent	M	1	6	19	26	8	1	9	32	15	3	1	0	3	6	50	2	5	27	14	12
	F	1	5	8	38	8	2	11	32	10	5	1	0	3	10	47	4	10	19	19	8

TABLE 1 (*continued*)

		Intelligent					Considerate					Independent					Inconsiderate				
		1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Disorderly	M	2	6	26	22	4	0	0	25	25	10	1	4	36	13	6	5	14	35	6	0
	F	0	1	32	20	7	0	1	23	19	17	0	0	41	15	4	9	11	33	7	0
Discourteous	M	0	4	31	21	4	0	0	2	20	38	0	4	42	10	4	13	31	11	5	0
	F	0	1	29	21	9	0	1	0	18	41	1	5	40	11	3	24	23	10	2	1
Enterprising	M	12	28	17	3	0	2	11	45	1	1	12	35	12	1	0	3	11	38	5	3
	F	14	34	10	2	0	2	11	37	8	2	22	27	11	0	0	8	10	32	8	2
Even-tempered	M	3	5	41	8	3	9	40	9	2	0	3	10	33	13	1	2	2	15	37	4
	F	3	14	40	3	0	18	31	9	2	0	1	13	37	7	2	3	3	12	35	7
Efficient	M	12	26	20	2	0	2	14	43	1	0	9	22	26	3	0	2	7	41	9	1
	F	14	36	10	0	0	5	14	40	1	0	13	32	12	3	0	2	7	36	13	2
Energetic	M	9	22	29	0	0	3	14	43	0	0	14	32	13	1	0	3	16	32	9	0
	F	10	20	27	2	1	5	13	39	3	0	23	27	10	0	0	6	12	37	3	2
Enthusiastic	M	10	17	30	3	0	7	23	29	0	1	11	22	22	5	0	2	8	30	18	2
	F	8	22	28	0	2	7	23	28	2	0	17	22	19	2	0	2	6	42	8	2
Conscientious	M	9	32	18	1	0	14	26	19	0	1	12	30	17	1	0	2	6	29	19	4
	F	16	27	16	1	0	20	25	14	1	0	12	25	19	3	1	1	6	19	28	6
Friendly	M	0	15	41	4	0	31	24	3	1	1	3	12	37	8	0	1	4	26	24	5
	F	7	12	40	1	0	33	23	3	1	0	3	20	33	4	0	2	4	33	17	4
Honest	M	2	29	27	0	2	13	30	17	0	0	10	17	29	4	0	3	3	46	8	0
	F	8	24	28	0	0	17	25	17	1	0	8	14	37	1	0	4	4	40	11	1
Hostile	M	1	1	37	18	3	0	0	5	25	30	0	13	35	8	4	7	17	27	8	1
	F	0	2	37	13	8	1	1	2	24	32	2	7	40	9	2	7	28	17	4	4
Intelligent	M	46	11	3	0	0	3	16	41	0	0	10	21	28	1	0	2	3	51	4	0
	F	51	7	2	0	0	4	12	42	2	0	8	26	26	0	0	1	6	47	5	1
Irresponsible	M	0	2	28	25	5	0	0	17	26	17	0	5	25	24	6	4	24	26	6	0
	F	0	2	14	34	10	0	1	7	35	17	0	5	15	27	13	11	25	19	3	2
Hypocritical	M	2	7	30	19	2	0	0	18	24	18	1	9	35	7	8	11	17	26	5	1
	F	0	6	27	18	9	1	3	13	24	19	1	7	30	12	10	16	21	16	2	5

*The numbers 1, 2, 3, 4, and 5 at the top of each column stand for the following responses: 1, "very often are"; 2, "tend to be"; 3, "may or may not be"; 4, "tend not to be"; 5, "seldom are." M, male; F, female.

TABLE 1 *(continued)*

		Intelligent					Considerate					Independent					Inconsiderate				
		1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Humorless	M	1	2	31	21	5	0	3	25	23	9	1	6	31	12	10	0	12	40	7	1
	F	3	1	28	21	7	0	1	28	22	9	0	2	42	10	6	4	12	36	6	2
Imaginative	M	15	34	9	1	1	6	14	38	2	0	15	27	16	2	0	4	6	33	14	3
	F	19	23	17	1	0	7	16	35	2	0	21	24	12	3	0	0	4	40	12	4
Impulsive	M	2	16	31	10	1	1	6	30	20	3	8	22	20	9	1	10	25	20	4	1
	F	0	7	32	18	3	1	10	28	15	6	12	18	22	6	2	14	20	21	3	2
Indolent	M	1	3	30	22	4	0	0	27	27	6	0	2	29	22	7	3	13	40	4	0
	F	0	1	24	23	12	0	2	17	27	14	0	9	20	19	12	9	10	33	1	1
Irritable	M	2	9	43	6	0	1	0	6	42	11	4	16	33	7	0	8	32	16	4	0
	F	2	5	44	6	3	0	2	6	31	21	1	11	34	13	1	17	26	14	2	1
Intellectual	M	10	30	19	0	1	0	5	51	3	1	6	23	28	2	1	5	4	45	6	0
	F	13	25	21	1	0	2	6	46	5	1	4	14	39	2	1	2	5	48	3	2
Impractical	M	2	7	31	17	3	0	2	36	14	8	1	5	32	18	4	3	9	39	8	1
	F	0	1	26	25	8	0	4	32	18	6	2	2	22	28	6	9	10	28	8	5
Independent	M	13	24	23	0	0	2	8	35	14	1	50	6	3	0	1	17	25	13	4	1
	F	20	27	12	1	0	3	6	35	14	2	49	6	2	1	2	21	14	17	4	4
Inefficient	M	0	4	19	33	4	0	1	40	16	3	1	2	28	25	4	3	7	47	3	0
	F	0	1	15	33	11	0	0	32	20	8	0	2	17	32	9	4	9	38	8	1
Inconsiderate	M	1	7	37	14	1	1	0	1	9	49	2	11	35	10	2	51	5	0	1	3
	F	0	2	34	18	6	0	0	2	5	53	1	17	32	6	4	54	3	2	0	1
Loquacious	M	4	13	38	5	0	1	5	36	16	2	0	15	35	9	1	4	15	37	3	1
	F	2	10	43	5	0	1	4	35	20	0	3	14	35	4	4	11	12	31	4	2
Neat	M	1	10	43	5	1	4	21	34	1	0	4	7	40	9	0	2	1	47	8	2
	F	4	11	42	3	0	9	19	31	1	0	5	13	39	3	0	2	3	37	17	1
Non-Intellectual	M	1	2	13	31	13	0	3	50	5	2	1	3	30	20	6	3	6	40	9	2
	F	1	2	18	22	17	1	5	47	4	3	0	2	36	13	9	2	3	50	2	3
Practical	M	1	24	29	5	1	3	17	38	2	0	5	26	23	6	0	5	7	36	8	4
	F	6	22	28	4	0	6	20	30	4	0	12	26	19	3	0	9	6	31	10	4
Modest	M	0	13	39	8	0	6	29	23	2	0	5	7	26	20	2	2	1	15	26	16
	F	2	15	37	6	0	11	31	16	2	0	3	4	30	21	2	0	1	20	25	14

TABLE 1 (*continued*)

		Intelligent					Considerate					Independent					Inconsiderate				
		1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Placid	M	3	5	36	11	5	1	15	33	10	1	2	13	19	18	8	0	5	22	24	9
	F	3	6	31	16	4	7	13	30	9	1	2	13	22	15	8	0	3	24	26	7
Unintelligent	M	0	1	3	10	46	0	0	36	16	8	1	1	23	28	7	1	2	46	10	1
	F	0	0	2	6	52	1	1	38	12	8	0	0	21	18	21	1	2	50	5	2
Reserved	M	5	18	34	3	0	2	16	33	9	0	5	20	21	10	4	3	4	24	21	8
	F	5	9	41	4	1	5	12	30	12	1	6	6	24	20	4	5	1	31	17	6
Reliable	M	11	25	21	3	0	22	22	16	0	0	11	23	23	3	0	2	2	23	24	9
	F	10	35	15	0	0	23	30	6	1	0	10	26	22	2	0	1	2	15	26	16
Sympathetic	M	3	23	24	9	1	32	26	2	0	0	4	10	27	19	0	1	2	8	30	19
	F	5	17	34	4	0	36	21	3	0	0	5	7	36	10	2	1	1	5	31	22
Responsible	M	11	26	21	2	0	14	33	13	0	0	16	17	25	2	0	3	2	31	16	8
	F	15	33	11	1	0	25	26	8	1	0	20	22	16	2	0	1	1	22	27	9
Sociable	M	4	14	37	3	2	13	37	9	1	0	3	9	33	14	1	3	5	25	21	6
	F	3	15	40	2	0	23	23	12	2	0	8	18	24	10	0	3	4	27	20	6
Sincere	M	3	25	29	1	2	21	29	10	0	0	9	20	29	2	0	5	7	21	20	7
	F	7	17	36	0	0	34	17	8	1	0	10	16	29	4	1	1	4	22	19	14
Submissive	M	0	2	18	27	13	1	13	34	9	3	0	2	7	21	30	0	2	9	35	14
	F	0	3	16	31	10	4	20	23	11	2	0	0	8	14	38	0	2	14	27	17
Sad	M	2	3	43	8	4	0	0	30	20	10	1	7	40	10	2	0	8	31	14	7
	F	3	2	44	10	1	0	2	32	18	8	1	2	37	16	4	3	5	42	5	5
Taciturn	M	0	6	41	12	1	2	5	36	15	2	2	7	26	20	5	2	4	36	14	4
	F	1	5	38	10	6	4	4	27	19	6	1	4	29	16	10	5	9	36	7	3
Unenterprising	M	2	2	16	30	10	0	3	36	20	1	3	6	10	25	16	5	6	36	10	3
	F	0	2	13	28	17	0	6	36	12	6	1	2	7	21	29	4	7	31	12	6
Warm	M	2	5	47	3	3	23	32	4	0	1	3	4	33	16	4	2	3	12	28	15
	F	2	13	40	5	0	39	19	0	2	0	3	7	30	19	1	0	0	18	24	18
Witty	M	7	24	27	2	0	1	12	40	6	1	5	15	39	1	0	5	12	30	10	3
	F	9	13	37	0	1	1	10	42	6	1	4	10	42	3	1	7	5	40	5	3
Unimaginative	M	2	0	12	29	17	1	1	36	13	9	0	4	21	22	13	0	15	31	9	5
	F	0	1	19	28	12	1	4	32	15	8	0	4	16	20	20	4	10	42	2	2

2. *Specific Trait Inferences*

Granted the overall agreement in specific inferences between men and women, are there any discernible differences in trait meaning that, because of their patterning, could be regarded as meaningful? We further examined all *list* traits where the number of definite inferences made by men and women differed by 16 per cent or more (10 out of 60 subjects), as shown in Table 1.

TABLE 2
DEFINITE INFERENCES

	Men			Women		
Given trait	*N*	Mean	*SD*	*N*	Mean	*SD*
INTELLIGENT	60	27.5	12.8	60	28.4	15.5
CONSIDERATE	60	30.7	12.9	60	34.0	11.9
INDEPENDENT	60	28.5	14.7	60	29.6	11.5
INCONSIDERATE	60	27.2	13.6	60	28.2	11.9
Totals	240	28.5	27.0	240	30.2	25.6

TABLE 3
EXTREME DEFINITE INFERENCES

	Men			Women		
Given trait	*N*	Mean	*SD*	*N*	Mean	*SD*
INTELLIGENT	60	6.4	6.8	60	7.9	8.0
CONSIDERATE	60	10.3	7.8	60	13.6	9.2
INDEPENDENT	60	7.3	7.9	60	9.8	7.8
INCONSIDERATE	60	6.8	8.2	60	9.5	7.9
Totals	240	7.7	15.4	240	10.2	16.4

Women, more than men, expect an INTELLIGENT person to be *efficient, responsible,* and *independent.* They also, more than men, expect an INTELLIGENT person not to be *unreliable, dependent,* and *irresponsible.* Men, on the other hand, infer the traits *sympathetic* and *witty* with greater frequency than do women. We may gather that our college women assign, more readily than men, valuable qualities to INTELLIGENT.

As for CONSIDERATE, women more frequently than men make positive inferences to *reliable* and *submissive* and negative inferences to *aggressive, unreliable,* and *indolent.* Women appear to deny, more definitely, negative connotations in this quality in others, although it is of interest to note that women regard CONSIDERATE people as more likely to be *submissive.* Is CONSIDERATE a "feminine" trait for women?

Turning to the *given* trait INDEPENDENT the following differences emerged. Men more frequently than women infer certain intellectual characteristics—an INDEPENDENT person is *intellectual* but *not humorless.*

Women perhaps are less willing to admit defects in INDEPENDENT people. They infer with greater certainty the trait *efficient*. They also expect an INDEPENDENT person not to be *irresponsible, impractical,* or *inefficient*. Independence, presumably, is an issue of more active concern to college women than to college men.

As for INCONSIDERATE, women are more critical. For them, INCONSIDERATE means more definitely than for men, among other things, *impractical* and *hypocritical*. For women, more than men, INCONSIDERATE people are not *neat*, not *responsible*, and not *conscientious*. Men appear to have a more "ambivalent" connotation of this trait than women. They also infer more often that INCONSIDERATE persons tend not to be *unimaginative*, or *non-intellectual*, or *enthusiastic*, but they are *witty*. Perhaps men excuse INCONSIDERATEness where it serves intellectual and therefore "good" values.

In sum, for women subjects, the preponderant emphasis of definite inferences is on such qualities as *efficiency, responsibility,* and *practicality*. For men, intellectual qualities emerge more clearly in their impressions; *humor, wit, imagination*. *Enthusiasm* also stands out for men. We may also speculate further that men less consistently than women place an all-embracing positive or negative value on a trait but rather are more likely to regard personal qualities as independent of one another.

E. Discussion and Summary

We have compared the inferences men and women make about a person's traits on the basis of a given single trait (INTELLIGENT, or CONSIDERATE, or INDEPENDENT, or INCONSIDERATE).

Men and women are very similar in the traits they infer from a *given* trait. However, women as a group tend to be more inclined than men to make positive or negative inferences of an extreme type. While the sex difference in this respect is slight, it is consistent over several independent groups of men and women, and for four different *given* traits. It is particularly true when the *given* trait is CONSIDERATE.

It is impossible to tell from the present data whether this result refers to differences in the implicit personality "theory" of men and women, and whether women actually behave more impulsively than men. Perhaps this is one of the things we mean when we say that women are especially "intuitive"—they are readier to entertain more definite extreme hypotheses on the basis of limited information. The difference in inference readiness and in extreme judgments between men and women, however, is to be considered together with the finding that education and social sophistication are not unrelated to extremity of judgment (3). It may simply reflect a response-set phenomenon, that would have been observed irrespective of the *content* of the rating scale used. As Osgood and Tannenbaum (2) have observed, extreme judgments are characteristic of "more emotionally oriented" individuals, such as women are supposed to be.

While men and women are also very similar in their connotations of the four given traits, some differences in trait meanings can be found that fall into a suggestive pattern. Women infer more often than men traits denoting such qualities as responsibility and efficiency. Men infer more readily characteristics related to intelligence and humor. The difference may be sex determined. However, it might also mean that the men and the women subjects may think of the "people" in the instructions as men and women respectively, thus yielding slightly different results.

In sum, if we give credence to these data, the implicit theory of personality of men and women is very similar. Yet women are slightly more ready than men to come to conclusions from information available about a person. Another way of looking at this would be that women conceive of personality as more definitely structured than do men. Other explanations are, however, conceivable. Also, the two sexes differ slightly in their connotations of the given traits.

When the specific content of impressions is considered, we find a high degree of similarity between men and women. There are indications of differential emphasis in trait inference. Women would appear to entertain a more consistently positive or negative connotation of a trait. Men seem less ready to commit themsevles, less "moralistic," more willing to acknowledge independence among personal traits.

REFERENCES

1. BRUNER, J. S., SHAPIRO, D., & TAGIURI, R. The meaning of traits in isolation and in combination. *In* R. Tagiuri and L. Petrullo (*Eds.*), *Person Perception and Interpersonal Behavior*. Stanford, Calif.: Stanford Univ. Press, 1958.
2. OSGOOD, C., & TANNENBAUM, P. H. The principle of congruity in the prediction of attitude change. *Psychol. Rev.*, 1955, **62**, 42-55.
3. SHAPIRO, D., TAGIURI, R., & BRUNER, J. S. Inferences in impression formation. (Unpublished report.) Laboratory of Social Relations, Harvard University, 1955.

Department of Psychiatry
Harvard Medical School
74 Fenwood Road
Boston 15, Massachusetts

Graduate School of Business Administration
Soldiers' Field Road
Boston 63, Massachusetts

Published as a separate and in *The Journal of Psychology*, 1959, **47**, 137-160.

PRE-RETIRANT TEST SCORES IN VARIOUS HOBBY GROUPS*

The American Life Foundation, Watkins Glen, New York

G. L. FREEMAN[1]

Several years ago the writer put himself under contract to write a book entitled *How Never to Retire.* But while gathering material on all types of people who had kept alert, useful, and productive after reaching the age of 65, he ran into an even more basic challenge—the *interests, attitudes, and problems of the pre-retirant.* He found to his dismay that surveys of retirement adjustment had grossly neglected high level professional workers and that their basic requirement was intelligent guidance in the years preceding a forced 'retirement' through social security policy. The need for research in this area was further emphasized when two advisees to Foundation Grants said: "Dr. Freeman, you retired from the psychology laboratory at 40 and built a new life around your avocation; how about turning aside at 50 from your writing-publishing career and undertake to design a counselling instrument for use with people who likewise desire a continued usefulness in the years ahead?" The result of this double-barreled challenge (need and opportunity) started a new Foundation to be devoted to retirement guidance counsel and a research program of which this is the third published report. The first report (3) began with the statement that "no preamble is required to establish the need for helping mature individuals plan for better adjustment by pre-retirement counselling," then proceeded with construction of the Interest-Values-Preference Test (IVP). Both there and in the validation report (1) which followed, it became fairly obvious that incidental learning and training in pre-retirant avocations was an important key to successful adjustment in later maturity. Again and again a productive worker of 50 was found to face no artificial age-of-retirement hazard when he already looked forward to work with some new or related skill which he now practiced avocationally. On the other hand, workers who reported no time for interest in anything beyond a very limiting vocation either looked on the arbitrary retirement age with dread—"I know I'll rot" or sought for pleasure in aimless ease—"I'll just sit out the rest of life in the sun."

---

*Received in the Editorial Office on November 29, 1958, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] USPHS Research Grant M-1034. The author expresses his thanks to Drs. P. C. and O. Smith of nearby Cornell University for statistical advice with reference to this report.

Undoubtedly there are some high level people whose mental adjustment in retirement is suited to withstand the deadening monotony and lack of challenge found in pointless quiet days. Not every man has a high energy level or can maintain at least its pretense through the ravages of failing health. But to those remaining purposive to the end, a dedication or hobby which defies time is vitally important. This present report documents the rôle of hobby interest as pre-retirant adjustment aids to healthful living in the years which lie ahead.

The prime objective here is to check interest-values-preference test scores (hereinafter referred to as *IVP*) against criterion data on pre-retirants with and without specifically confirmed hobby interests. The following points are discussed: (*a*) rationale of research procedure, (*b*) nature of the population samples utilized, (*c*) criterion measures used, (*d*) interests validity measures thereby obtained, (*e*) hobby vs. non-hobby avocational patterns in pre-retirants, (*f*) life values and personality factors in relation to interests, aptitudes, and training, (*g*) testing mental adjustment of pre-retirants, (*h*) summary, conclusions, and prospects for further research.

A. Rationale of Research Procedure

Two earlier reports have fully discussed *IVP* test construction, item selection, and analysis for improved forced-choice pairings. The present research concentrates on giving the revised test form to randomly sampled hobby groups of varied general educational (training) status and psychological adjustment in various parts of the country, then obtaining as much information as possible on the degree of individual interest and skill in that hobby. These so-called "criterion values" were obtained in three ways: (*a*) by direct peer-count from a hobbyist's own associates, (*b*) by self-appraisal questionnaire rating, and (*c*) in a few cases—by answers made by the subject to open-ended questions based on an idea developed from Sargent's Insight (projective personality) Test (5). In distinction to the structured private interviews which were basic in our first two reports, no such interviewing was conducted here. Procedure was as follows: (*a*) investigator appeared before an assembled group supposedly dedicated to one of the interest scales tested by *IVP*, spoke briefly of his own interest in their particular hobby and on life adjustment in general; (*b*) asked only those between the ages of 45 and 65 to take the 30-minute test and its accompanying questionnaire; (*c*) then in a hold-over group asked subjects to make a peer-count rating check of each other; with a few of the more interested also asked to work on a self-insight test. Since the *IVP* test is itself copyrighted by American Life Foundation and in present stage of development not available for use by outside agencies, here we show only the self-appraisal questionnaire and the Peer-Count Rating method for appraising related topics (Table *A*).

It can be appreciated that questionnaire and peer-count rating terms are subject to varied interpretation; but though often tempted to read minds, at no time did the investigator enter his own appraisal of an individual subject either in regard to depth of hobby interest or normalcy of psychological adjustment. The whole procedure was—so far as possible—carried out by an informal "have fun" group.

> Arriving with his armful of *IVP* tests, the investigator would try to enter into the group spirit and assume virtuosity. For example, opening with a "Barbershoppers" group in Chicago, he sat down to the piano, then rattled off his one and only show piece ("You Were Meant For Me") and got everyone singing; with the *TV* bowlers at Louisville he said, "Now when *I* used to follow this sport, all you had to do to get a 'Strike' was close your eyes and Wham!" wherewith the ball not even making the alley—a friendly "razzing" brought all the others up to tell the "Doctor" how really good *they* were. All such "horseplay" formed an essential device for getting adults (usually wary of taking tests and under no obligation to take time out from their play) to coöperate fully, telling not only on themselves but on their hobby associates as well. The informal "gab-fest" following test-taking also gave the investigator-researcher a perfect opportunity to sound out certain ideas gained through years of informal discussion with preretirant hobbyists. While such give-and-take had no bearing on test results, it was useful in pointing up problems requiring further research. By such means the investigator made his way from one "dedicated" special hobby group to another usually ending on this note: "I trust this was as much fun for you. You know it isn't often that a psychologist gets let into the inner sanctum of a pet-fanciers' fraternity (or whatever it was). Practically all of you make your living in some other field, but now that I've revealed *myself* vocationally, you do know that hobbywise I'm still one of the brothers. How many agree that here one can be completely himself? Here no one asks how educationed you are, how much money you have, whether your ancestors came here on the Mayflower or not. . . . For many this club then is their secret "Walter Mitty." Furthermore some are already building here a second vocation—something that will keep one young, active, and interested when the so-called retirement years catch up with one. I do have, however, just one caution from hobby group meetings like this. While a bank president and a truck driver may be "buddy-buddy" in this hobby group, one doesn't go into the bank tomorrow and ask this brother to give him an unsecured loan. A purely avocational personality interest can hardly go that far."

Naturally, the investigator would have liked to know more about the vocation and educational background of his volunteer subjects; but by keeping

TABLE A
SELF-RATING QUESTIONNAIRE

1. For each of the following activity areas place a check mark indicating some position on this scale which will stand for your own estimate of your present knowledge, skill or competence in that activity.

. . . .

Extremely Incompetent	Art and Music	Extremely Competent
Extremely Incompetent	Collecting and History	Extremely Competent
Extremely Incompetent	Intellectual Games & Mental Skills	Extremely Competent
Extremely Incompetent	Sports & Athletics (Part. or Spect.)	Extremely Competent
Extremely Incompetent	Social (Private Contact, Helping Others)	Extremely Competent
Extremely Incompetent	Nature, Pets or Plants	Extremely Competent
Extremely Incompetent	Science, Mechanics, Inventions	Extremely Competent

2. For each of following areas indicate by position check your degree of interest.

. . . .

Extremely Uninterested	Art and Music	Extremely Interested
Extremely Uninterested	Collecting and History	Extremely Interested
Extremely Uninterested	Intellectual Games & Mental Skills	Extremely Interested
Extremely Uninterested	Sports & Athletics (Part. or Spect.)	Extremely Interested
Extremely Uninterested	Social (Private Contact, Helping Others)	Extremely Interested
Extremely Uninterested	Nature, Pets or Plants	Extremely Interested
Extremely Uninterested	Science, Mechanics, Inventions	Extremely Interested

3a. Circle one of above interest areas presumably bringing this group together.

3b. How long have you participated in this kind of activity? Less than 1 year ___ More than 1 but less than 5 ___ More than 5 but less than 10 ___ More than 10 ___

4a. Indicate with a check your own principal hobby interest only if different from the one held for this group. 1 ___ 2 ___ 3 ___ 4 ___ 5 ___ 6 ___ 7 ___

4b. How long have you been interested in this other particular type of activity? Less than 1 year ___ More than 1, less than 5 ___ More than 5, less than 10 — More than 10 —

5a. Indicate in rank order any you formerly enjoyed but which now have somewhat lost their appeal. 1— 2— 3 — 4 — 6 —

5b. What was the longest period of time during which you actively pursued the most persistent of these six interests which you have ranked above. ___________

6. In your opinion, how do most people go about picking a hobby? Consider the following causes and place a rank of 1 before that which seems most important, 2 for the next, etc., on down to 6 as least important.

___ They have an aptitude and accordingly choose hobby they can do well.

___ They enjoy company of the kinds of people who are engaged in these hobbies and join them for companionship.

___ They choose hobby they can best afford in terms of time, money or health.

___ They choose hobbies which are suited to their social status in the community.

___ Any other factors in which a hobby contributes to a hapyy life.

7. What life-factors do you consider most essential to well-adjusted retirement? Place in rank order: Amusement ___ Purpose ___ Self-interest ___ Social Service — Conformity ___ Non-Conformity ___

8. Relation of Personality Factors important in successful retirement planning: Place in rank order: Energy ___ Emotional Control ___ Objectivity ___ Sociability ___ Ascendancy ___

TABLE B

LEADER DISCUSSION, QUESTIONNAIRE FORM, AND RATING METHODS FOR HOBBY GROUP PARTICIPANTS

Leader discussion:

"All of us are interested in this particular field as indicated by our attendance tonight. For instance, some are very interested and devoted to this area. Some may be here more because of mutual friends or just for something to do that takes us out of ourselves. A few even may be here because it's a thing to do with the Jones or a duty to a more dedicated wife or husband. . . . Now that *you* have taken the interests test and put down your general reactions on the questionnaire parts listed on back of the answer sheet, I wonder if you'd mind rating each other. Here is how that is done. In this small hold-over group, as I call a given member's number (like 1) he stands and all the rest of you thereby individually rate him to the best of your knowledge. Thus if you think subject No. 1 is best bowler (etc.) in this group of ten, write his number (1) in the + or best column; if the worst, write his number (1) in the — or worst column, or if somewhere in between, place his number along the B-W continuum. Now when No. 2 stands up for rating, if you think he is just as good or bad as (1), put his number in the same circle, etc. The form shown below calls for eight different ratings on each subject, so we go through all ratings on the list before changing to another subject. Do not rate yourself—however, while all the others are judging you, try to take the point of view of your spouse or best friend and rate as she or he would rate you. Above all, keep these ratings secret—it's just a game anyway; —and since this investigator doesn't even know your names, your rating is safe with him. You know a peer-count is quite 'Army.' Any questions? Let's go!"

Example of individual Peer-Count ratings made for a 10-subject group: Note: First subject (No. 1) was rated near top on Item 1*A* and subject No. 3 near bottom and subject No. 5 in between.

Interest Item:	Rank:	Best Near Top	— In Between —	Worst Near Bottom
1a. Good at this hobby and devoted to its pursuit		1	5	3

Life Values Items:	Rank:	*Personality Items*:	Rank:
1. For him life is more purposeful than just for amusement alone		1. Is he high or low on energy and activity	
2. He is more conformist than non-conformist (individualist)		2. Is he high or low on ascendancy	
3. He is more selfish than social service minded		3. Does he have good or poor emotional control	
		4. Does he have high or low objectivity (sees self as others see him)	

obvious "testing" in the background as much as possible, a sample of at least 20 dedicated hobbyists in each of five interest scale areas (plus 40 more with no hobby affiliation) were carried through the complicated forced-choice *IVP* test and the difficulties of attendant criterion checks. More than that could not be handled under the research procedures thereby developed.

B. NATURE OF THE SAMPLE

Whereas previous work with the *IVP* test had dealt exclusively with executive level pre-retirants, here it was desired to obtain information on a more general population sample vocationally. However, if we were to have any adequate criterion checks, some selection was essential. Acting on the presumption that the interests scale is the most valid part of *IVP*, we made a crucial test on pre-retirant subjects of varied vocational and educational background *known to be associated in a definite hobby-interest group*. This was made possible through the coöperation of various hobby clubs in the East and Middle West; the investigator journeyed to their places of meeting and administered the *IVP* test to 10 or more people sitting as a group of artists, nature lovers, or "what have you." The only stipulation held was the age-range of 45-65; sex, education, or other status cues were ignored and the entire hobby group encouraged to complete test-taking at that single evening session. Coöperation, as might be expected, was not uniformly good. Mature individuals under no obligation to take a test often had to be cajoled into even starting to fill in the forms, while others participated to their full enjoyment throughout the session. It was found that our investigator could handle a group of 10 with more worthwhile results than when group testing involved larger combines. And soon this procedural pattern came to be found most practical with the majority of situations:

1. In a given geographical location we collected 10 cases from a group covering one aspect of a general hobby interest area such as sports (e.g., 10 men meeting as Louisville, Kentucky, tournament bowlers); then we obtained 10 added cases for a group covering another aspect of the same general (sports) area but from some other geographical location (e.g., 10 members of a Troy, New York Sportsmen's Fin, Fur, and Feathers Club).

2. We did likewise with each of the seven general interest scales covered by the *IVP* test; so that the overall total sample would carry 20 *sports* hobbyists, 20 *art and/or music* hobbyists, 20 *collector-history* enthusiasts, 20 who belonged to *pet or nature* clubs, 20 *mechanics-science* hobbyists, 20 pre-retirants whose avocational center was *mental games and skills,* and 20 in *social service* hobby groups. Unfortunately, in the general interest areas of *mental skills* and *social service,* we were unable to find any organized groups where either the potential subjects for the experiment were sufficiently interested or the majority were sufficiently dedicated to that area to say, "yes, this (bridge playing or good cause X) is my major hobby." Therefore, we compromised with only five dedicated major interest groups (*A,* Art-Music; *B,* Collecting-History; *D,* Nature-Pets; *E,* Mechanics-Science; *F,* Sports).

3. In place of confirmed hobbyists in Areas *C* (Mental Skills) and *G* (Social Service), we lumped together a random 20 cases composed of people who said that while they "rather liked" one or the other of these diversions,

it was not a major hobby with them. In fact, this group had no special hobby so we have called it the *UC* or "uncommitted class."

4. Finally as a fillip to working with dedicated or committed hobby groups, we dug around for another 20 cases involving men who (while belonging to one of several business clubs) said they "had no hobbies or basic interest beyond making money." This we have called the *QG* or "Questionable Group."

5. It would have been a neat experimental design if (with the 10-10 break between two different hobby groups in the same 20 case general interest category) we could have matched a predominantly spectator group with a predominantly participant group. But we had enough trouble gaining any loyal group coöperation to hold out for such consistencies. Nor could we follow our original intention of having the same number of men and women in each 20-case group. (If more cases were available, those cases were selected at random). What we came out with, design-wise, is shown below by the following category-breaks.

A. Art and Music Interest Groups

Ten who belonged to a music-harmonizers group (mostly men), 10 members of an art league (mostly women).

B. Collecting-History Interest Groups

Ten who belong to The Early American Glass Club (mostly women), 10 members of a State History Society Group (mostly men).

UC. Mental and/or Social Skills Interests, but yet *Uncommitted* Group

Twenty men and women who either did play bridge (etc.) or did go to meetings connected with social service projects, but who said this was scarcely a major hobby.

D. Nature and Pets Interest Groups

Ten dog breeders also following Horse Shows (mostly men), 10 members of a garden club (mostly women).

E. Mechanics and Science Interest Groups

Ten men and women model railroaders, 10 men and women in a science study club.

F. Sports Interest Groups

Ten tournament bowlers (mostly men), 10 members of a sportsman's club (mostly men).

G. Questionable Interests Group (20 non-hobby men)

Analysis: The question that immediately suggests itself is whether test profiles are distinctive for these differently selected interest groups; also the significance of other parts of the *IVP*. But before sketching such profiles, we need to indicate further aspects of the experimental sample itself: Specifically, what problems arise in getting questionnaire replies and peer-counts from volunteer adult subjects when the intent is to commit them concerning whether they think they are good at this hobby, how dedicated they are and the presumed relation to life values and personality adjustment. Not all such questions received replies—only enough to be suggestive. In general, a hobbyist seldom hesitated to commit himself as to how "good" he thought he was or his degree of interest compared to other activity areas. Other parts of the questionnaire were frequently left blank.

On the matter of getting peer-counts, even more difficulty was encountered. Our greatest success in eliciting coöperation came only near the end of data-collection. Each man was invited to take his turn in *standing up* before the group while colleagues made a forced-distribution rank on each, with reference to eight pointed queries. Such a peer-count as this could then be put against actual test scores on three different parts of the *IVP* test. It is regrettable that such data could not be made more complete, and therefore are of suggestive value only.

The number of cases with even nearly complete projective (self-insight) test-data was so very limited it need not be discussed here. Suffice it to say that had we worked with paid or institutionalized subjects in the age range desired, 45-65, some further insight into the personality adjustment of preretirants might well have been acquired. The disadvantage of such a sample is equally obvious; whereas here we had men and women still working at one of many different vocations, the institutionalized case is by definition a person who has already withdrawn from the world of work; hence is *not* a preretirant.

C. Criterion Measures Used

This is always the big question in any test study of validation rationale. Since our first problem was to test the validity of activity preferences in preretirants, our chief criterion measures were the selected hobby groups in each of the five general interest areas (art-music, collecting-history, nature-pets, science-mechanics, and sports). Such average test profiles (20 cases each with no distinction as to sex, education, or vocation) indicate that each area's hobby group scores high on the appropriate section of the interests test. Note also that criterion hobby groups tend to overlap in degree of related interest. For instance, "art" and "collector" groups both have less interest in participant sports appeal than do the "nature" and "mechanics" interest groups. The overall means and medians obtained for these 100 "committed" hobbyist cases are shown in Table 1. This is apparently quite representative of scores obtained in the previous report where we did not know the hobby interest or degree of dedication to the subject.

Since the question will be asked, we here state that comparison between profiles of similar sub-groups, i.e., bowling and outdoor sport, yielded no fruitful cues to differentiation; in fact, such differences in no cases were significant at the one per cent level; hence, we assume that "criterion-wise" hobby groups, taking the interest part of *IVP*, do not specifically pinpoint the different kinds of interest in the same general field nor the degree of spectator-participant activity thereof. Anyway, it would be asking miracles that such small sub-group samples as were used in this study could all produce interest measures specifically identified with a particular branch of the hobby. Far more important is the fact that regardless of sex, educational background, or location of the samples taken, overall average profiles for each hobby group used tend to highlight the general interest area we set out to test.

TABLE 1
OVER-ALL MEANS AND MEDIANS FOR RESPONSES MADE TO SEVEN DIFFERENT INTEREST SCALES BY FIVE DIFFERENT HOBBY GROUPS
(N = 100 Ave. 8.7)

Scale		Over-all means	Over-all medians
A	Art and Music	8	8.5
B	Collecting and History	10.6	9.5
C	Mental Skills	8	8
D	Nature and Pets	9.7	9.5
E	Mechanics and Science	8.4	7.5
F	Sports	8.6	7.5
G	Social Service	9.1	9

Since in the final analysis we could use only subjects present at a group meeting who said this was their major avocational interest, it will be seen that our selected sample is not exactly random. On the other hand, such criterion selection cuts the randomness factor only for the interests part of the *IVP* test; for remaining parts of the test study, sampling was as random as anyone could wish.

Identifying pre-retirants whose avocational interest and activity places them in a definite hobby group was only a small part of the *IVP* test's total criteria problem. What psychologically based factors other than mere physical attendance at such a group meeting ought to be studied for criterion purposes? Certainly not all members would be equally interested, proficient, or psychologically well-adjusted in a given avocational field. For comparison purposes we therefore obtained on individuals in each of the five hobby groups, forced-distribution measures both by self-ratings and by peer-count. Then for contrast purposes we also took two 20-case pre-retirant samples that *were not highly oriented avocationally*. The *UC* (or uncommitted "group") was not an organized hobby group at all; instead, the 20 cases were selected upon basis of their 45-65 age background and their statement of not being adverse to playing at "mental games" or "social service," two interest scales not covered by one of the five dedicated hobby groups. The *QG* (or questionable group) of 20 cases was similarly placed, except that here men were selected on the basis of an announced statement *they had no hobby*.

D. VALIDITY MEASURES OBTAINED FOR *IVP*'S INTEREST SCALES

This section of the report deals exclusively with the interest part of the *IVP*. Under the conditions set up for our criterion (committed hobbyists meeting in a special group), test results should correlate highly therewith. To check validity, we used that particular variation of non-parametric statistics (6) known as the Median Test, set at the one per cent level of significance. For this (*IVP* Test 1) we first determined the overall median score of a given interest sub-scale for subjects in all five dedicated hobby groups (total 100); then for each related scale we made up Chi Square tables showing number of cases above and below that median for its related hobby club (total 20 cases) and the number of cases for all unrelated hobby clubs above and below that median (total 80 cases). These Chi Square values are shown in Table 3, while various sub-scale scores appear as Table 2.

TABLE 2

AVERAGE SCORES FOR 20 SUBJECTS EACH IN A DIFFERENT HOBBY GROUP AS SHOWN BY OVER-ALL INTEREST TEST SCALES

	Scale	A	B	C	D	E	F	G
A	— Art-Music Hobbyists	14.8	9.5	8	8.8	5	7	9.6
B	— Collecting-Historical Hobbyists	8.9	17	7.8	8.6	5.9	5.2	9.4
C	— Mental Skills (no sample)							
D	— Nature, Pets Hobbyists	8.8	10.2	5.4	14+	8.4	8.2	8.4
E	— Mechanics-Science Hobbyists	8	7	10	8.6	15.2	6.6	8
F	— Sports Hobbyists	6.4	6.6	8.6	9	7.8	14	10
G	— Social Service (no sample)							
UC	— Uncommitted Hobby Group	8	7.9	9.2	10.6	8.2	8.4	10.2
QG	— "Definitely no hobby" Group	7	7.2	10.2	9.2	5.4	9.9	9.4

As can be shown by inspection of *IVP* interest Test 1 at the top of Table 3, cases forming a specific hobby group scored upon the associated sub-scale *above* the overall (100 case))median. Applying the *median test* to these Chi Square values, we find that in all cases the *IVP* interest scores are well above the level of significance required as set up for a "two-tailed" test.

Since each of five different hobby groups was above the overall median for its own scale, the interests part of the *IVP* does test the prime avocational interest of a dedicated hobby group.

As for the two scales involving no dedicated hobby group sample, results here are as should be expected. Lacking specifically designated hobby groups to test for *mental skills* and *social service* scales, we used the 20 *UC* cases for whatever this might show. Having tried to make sense out of individual interest profiles in the sample, we decided to treat the data only as *if* these responses to the "mental skills" scale represented an uncommitted hobby group at least not adverse to that activity field, then to treat the same *UC* group's responses to the "social service" scale as similarly placed in terms of interest. From this we went on to apply the "two-tailed" median test for indicated significance covering the two remaining *IVP* interests scales (*C*—Mental Skills and *G*—Social Service). To make such a non-parametric comparison we added scores of the 20 uncommitted cases in with the 100 confirmed hobbyists already studied and so found the overall medians for the mental skills and social service scales. Chi Square values for these 120 cases are indicated by the asterisks (**) in the *IVP* interest Test 1 at the top of Table 3. Following the method previously described, we came out with a p value of .375, well below the level of significance required.

This brings us to the average *IVP* scores of each hobby group as shown in Table 2. Note that while the *UC* and *QG* groups are not significantly higher than the committed hobby groups on any of the interest sub-scales, they do average slightly higher on the social service scale; also that the *QG* has its highest score in the mental skills area. When we come to a Chi Square test of significance, however, the p in both cases is above the five per cent level of significance. Thus neither an uncommitted (*UC*) group nor a non-hobby (*QG*) group would be sufficient for criterion study of a scaled interest test. We consider that our committed hobby groups adequately show that the interests part of *IVP* tests what it purports to do.

E. Hobby vs. Non-hobby Group Avocational Patterns in Pre-Retirants

Though our 40 case (*UC* and *QG*) non-hobby sample is not as large as dedicated groups cases, we shall here attempt an overall comparison of interests and background. What exactly would the median test of significance show if both committed and non-hobbyist interest scores were treated together; then how might these scores be related to educational background and incidental training?

IVP Section 2 of Table 3 shows the Chi Square values covering interests distribution of 40 non-hobbyists (20 *UC* and 20 *QG*) and 80 hobbyists (20 Art, 20 Collecting, 20 Nature, 20 Mechanics). When the number of hobby and non-hobby cases above and below the overall median for each interest sub-scale is set down and the median test of significance is run, we find that

TABLE 3

MEDIAN CHI SQUARE TESTS OF SIGNIFICANCE FOR ALL COMPARISONS OF *IVP* TEST SCORES

IVP Interests Test 1	Art — Music			Collecting-Hist.			Nature — Pets		
	Above M = 8.5	Below M = 8.5	T	Above M = 9.5	Below M = 9.5	T	Above M = 9.5	Below M = 9.5	T
Specific hobby	20	0	20	20	0	20	20	0	20
Other hobbies	30	50	80	30	50	80	32	48	80
p =			< .01			< .01			< .01
IVP Interests 2	M = 7	M = 7		M = 9.5	M = 9.5		M = 10	M = 10	
Hobby	62	18	80	52	28	80	38	42	80
Non-Hobby	11	29	40	8	32	40	20	20	40
p =			< .01			< .01			NS
IVP Values 3	Amusement			Purpose			Conformity		
	M = 4	M = 4		M = 4.8	M = 4.8		M = 4.5	M = 4.5	
Hobby	40	40	80	50	30	80	36	44	80
Non-Hobby	24	16	40	12	28	40	14	26	40
p =			NS			< .01			NS
IVP Personality 4	Energy			Emotionality			Objectivity		
	M = 5.5	M = 5.5		M = 7	M = 7		M = 5	M = 5	
Hobby	56	24	80	44	36	80	44	36	80
Non-Hobby	4	36	40	24	16	40	16	24	40
p =			< .01			NS			NS

T = Total.
M = Median.
NS = Not significant.

TABLE 3 (*continued*)

IVP Interests 1	Science-Mech.			Sports			Mental Skills**			Social Service**		
	Above M = 7.5	Below M = 7.5	T	Above M = 7.5	Below M = 7.5	T	Above M = 8	Below M = 8	T	Above M = 9	Below M = 9	T
Specific hobby	20	0	20	20	0	20	12	50	62	13	50	63
Other hobbies	31	49	80	30	50	80	8	50	58	7	50	57
p =			< .01			< .01			NS			NS
IVP Interests 2	Science-Mech.			Sports								
	M = 9	M = 9		M = 7	M = 7							
Hobby	34	46	80	32	48	80						
Non-Hobby	20	20	40	28	12	40						
p =			NS			< .01						
IVP Values 3	Non-Conformity			Selfish			Social					
	M = 3.5	M = 3.5		M = 5.5	M = 5.5		M = 5.5	M = 5.5				
Hobby	32	48	80	28	52	80	40	40	80			
Non-Hobby	27	13	40	32	8	40	14	26	40			
p =			< .01			< .01			NS			
IVP Personality 4	Sociability			Ascendancy								
	M = 6.5	M = 6.5		M = 6	M = 6							
Hobby	60	20	80	28	52	80						
Non-Hobby	26	14	40	29	11	40						
			< .01			< .01						

**This group under Hobby is actually Uncommitted.
T = Total.
M = Median.
NS = Not significant.

the non-hobbyists score significantly below the dedicated hobbyists for both art and collecting categories, but not significantly below the overall median scores for nature and mechanics categories. Only in the sports category are non-hobbyists significantly *above* the overall median score for that area; and this is of no significance either, since we did not utilize in our comparison sample the dedicated sports enthusiasts (least homogenous and least dedicated of all hobby groups studied).

As for the Chi Square median test of significance for the Values and Personality test scores (Table 3, Sections 3 and 4), we see some significant differences between hobby and non-hobby groups; these may, of course, be attributable to sampling errors.

With regard to differences in background between hobby and non-hobby cases, perhaps the most significant finding of this entire study is that hobbyists represent a higher degree of formal education than do the non-hobbyists in our sample ($p = < .01$). Since it was not in our experimental design to explore the *rôle of learning as both a broadener and developer of avocational interests,* we have little but questionnaire reports to go on. From that and informal observation of subjects, however, we should suspect that the more an individual educates himself formally or informally, the wider his understanding and sweep of interest; and the more he incidentally trains himself, the better off he is vocationally and avocationally.

Another approach to validation of the *IVP* avocational interest scales is through comparison with self-analysis questionnaires and peer-count ratings. Reference to the questionnaire form shown in Table *A* indicates how at least a three-point rating (high-median-low) was obtained on the relative competence, etc., reported by each subject for all avocations covered, including the hobby group to which he belonged. In the test of significance we have analyzed variance between questionnaire measures of alleged competence, etc. in a hobby and the test scores made upon such a part. Like the previous Chi Square test, the indication is that the 20 cases in a specific hobby group are rated more competent than the 80 others who were in some other hobby group. This is also borne out by comparisons based on peer-count, with significance above the .05 per cent level. Furthermore, correlations run between actual test scores and competence self-rating made thereon are highest for a related hobby group. Not only do hobbyists of a given interest pattern rate themselves and each other more competent there than in other areas, they are also more interested in their chosen field, have followed it on an average of five or more years and rank other interests below their chosen field. The table of r's, from an overall point of view, indicates that dedicated hobbyists have a competence rather matching degree of interest and length of time pursued. Since such self-analysis rating is borne out by peer-count, this would suggest that training (from continued practice in a field) may be a basic factor in hobby interest; in other words, interest breeds competence and it may be that competence holds interest! But with this the case, what then first creates or makes for an interest? Fortunately, we have a partial answer by correlating questionnaire and test results Regardless of which of five hobby groups is surveyed, the two chief factors given for an interest were—"enjoys company of those with similar likes" and "has an aptitude for competence in that field." The following percentages show how this rank-order (for an average 100 hobby cases) stands with regards to why a particular hobby is chosen: enjoy company, 41 per cent; have aptitude, 29 per cent; chosen in terms of money or health, 15 per cent; urged on by family and friends, 14 per cent; chosen as suited to status in community, one per cent. But when we partial-out individual cases in terms of a high rather than a flat interest profile, no significance whatever is seen.

Derivation of interest and its relation to vocation are big problems this study scarcely touched upon. Only when one takes the peer-count on emotional adjustment and compares it with self-analysis of why people choose hobbies and what makes them variable, does one find a nearly significant difference. For the correlation between peer-count on emotional control and self-adjustment score is .25 and between energy level score and rating is .30, while the rating of those placed as the in-between peer group and the objectivity score gives an *r* of .29. In other words, hobbyists and their peers agree that high working energy, good objectivity, and emotional control are factors in successful pre-retirement adjustment. Less surely they seem to show some agreement (peer and self-analysis) that good sociability and lower ascendancy help also. The critical test really comes, however, when value and personality test scores are compared for hobby and non-hobby groups.

F. Life Values and Personality Factors in Relation to Avocations and Aptitudes

Having shown an indication that those who belong to a given hobby group tend to be interested, competent, and emotionally attached thereto, it next behooves us to look at results obtained on the other two parts of the *IVP*—specifically, the life values and personality factors. Is there any evidence, for instance, that people in a dedicated hobby group have a significantly different set of life values or personality patterns than those who do not? Here we have attempted to study only a selected partial analysis to the data; that is, we have taken the 40 cases which had no special hobby (at least belonged to no defined hobby group) and made several Chi Square tests of significance; first asking whether there was more or less of purposiveness in life, more or less conformity(this for Values); then asking was there more or less sociability and energy (this for Personality). We thus have utilized only two items each of the Values and Personality parts of *IVP* test. The indication from an analysis of variance test is that the confirmed hobbyist (one already *centered* at this pre-retirement age 45-65) is still more the conformist than the non-conformist, more purposive than purely amusement-bent; also (unlike the non-hobbyist) he is more energetic and objective ascendance-wise and routinely selfish. These are not hard and fast findings, but suggestive only of more work.

TABLE 4
COMPARISON OF AVERAGE *IVP* VALUES AND PERSONALITY TEST SCORES FOR 120 CASES FORMING THREE GROUPS

Values score	Amuse.-Purp.		Risk.-Cau.		Conf.-N.C.		Self.-Soc.		Ethics
Group 1	4.0	3.7	2.3	3.1	4.7	3.1	4.8	5.5	4.9
Group 2	3.2	5.5	3.4	2.6	4.7	4.0	3.2	5.5	4.3
Group 3	6.1	3.1	1.3	1.8	3.2	3.9	6.9	4.9	4.8

Personality score	Activity	Emotionality	Objectivity	Sociability	Ascendancy
Group 1	6.1	7.1	5.0	7.5	4.6
Group 2	7.2	6.3	4.8	6.1	5.7
Group 3	4.0	7.5	4.4	7.0	7.1

Group 1 — Art-Collecting Hobbyists.
Group 2 — Nature-Mechanics Hobbyists.
Group 3 — Non-hobbyists (*UC*'s and *QG*'s).

Tables 4-8 show the average scores of both hobby and non-hobby groups on Values and Personality Parts of the *IVP* (120 cases in all). Notice that we separated the two 20-case studies in the arts and collecting Group 1 from the two 20-case studies in the nature and mechanics Group 2, this because we felt there might be a significant relation of "crossline" interest, life values, etc. Though none is especially apparent, it would seem that our Group 3 cases (non-hobby 20 *UC*'s and 20 *QG*'s) give different types of average profile. Of course, the real check here is the median test of significance as shown in Table 3. These Chi Square values of Table 3, Sections 2, 3, and 4, tell their own story not only for the Values and Personality parts of *IVP* test, but the related interest scale differences as well. The Values scores for the 40 non-hobbyists is not significantly different from the 80 hobbyists so far as amusement, conformity, and social service attitudes are concerned. On the other hand hobbyists score significantly less selfish than the non-hobbyists and also less significantly non-conformist. The non-hobbyists are less purposely motivated in life than the hobbyists; but this difference barely approaches significance in the present test results. The indication is, however, that non-hobbyists (in contrast to the dedicated hobbyists) are more selfish, more non-conformist, and less purposive than those found in dedicated hobby groups. Comparing these results with data obtained on questionnaire and peer-counts, we would have expected the amusement factor to be stronger in the non-hobbyist and the selfish and non-conformity factors to have been approximately the same for both non-hobbyists and hobbyists. The fact that significance *is* here present provides thought for later discussion and research.

People who are "happy" in their avocational retirement often seem also to have been happy in their old (vocational) job. If this is true, serenity becomes a base line for both vocation and avocation and the problem of shift to either related or different field of activity after 65 is a matter of timely "think-help" and sanely practiced life values before!

In addition to the Life Values scores, these final paragraphs are concerned with the personality part of the *IVP*. Here scores on each of the factor scales have first been checked for overall median, then the number of cases above and below such median have been separated into non-hobbyist and hobbyist groups. Applying the median test to the Chi Square values of Table 3, Section 4, we find that on three personality factors there is a significant difference between hobby and non-hobby cases. The hobby group tests significantly higher on general activity (energy) than the non-hobby group and is also less ascendant and more sociable than the non-hobbyist. This finding accords with correlations run between test scores and forced-distribution on related aspects in the questionnaire form ($r = .34$ activity, .26 sociability and .28 ascendancy) for hobbyists. As for the factor of emotional control the r's for the hobby group, while positive and greater than for non-hobbyists (.27 vs. .21) is not borne out as significant by the median test. In fact, both on emotional control and objectivity the low positive r's for both hobbyist and non-hobbyist show that these factors lack significance when studied in terms of test scores alone.

We are ready to conclude, therefore, that for the test data here analyzed, significant personality factors dealing with successful avocational adjustment of pre-retirant hobbyists was limited to fair objectivity wtih low ascendancy and high energy factors. Here we have also shown that both personality and life-value factors are related to avocational aptitudes,—with a strong suspicion that incidental training plays a major rôle along with personality and value factors in the choice of satisfying and well-performed hobby activity by pre-retirants. Other and more subtle interrelations to the psychology of pre-retirant adjustment are discussed in the next section.

G. TESTING MENTAL ADJUSTMENT OF PRE-RETIRANTS

Quality or type of mental adjustment met in pre-retirant hobbyists was never a major question in this research. Nonetheless, apparent differences seen in casual observations of subjects taking the *IVP* suggested that some slight introduction of quantitative measures to this field be attempted. What every retirement counsellor would like is the presently quite impossible objective screening test of obvious mental unbalance—a guide enabling the counsellor to refer such cases to proper clinical authority as beyond his purview. Let no one think that a dedicated hobbyist nearing retirement is always psychologically well adjusted. Some "queer" people inhabit these ranks as well as many who are "salt of the earth." Often the pretenses and dream-world compensations of an oldster serve to *ego homeostate* the personality and to *maintain* him as a still useful member of society. Here we need not be concerned with such harmless aberration, rather look for maladjustment trends that unless corrected can push the oldster and his family towards a serious mental health problem.

Casting about for 'first straws' to this area of retirement counselling, we hit upon a much modified version of Sargent's (projective) Insight Test. Using only five "armatures" of our own devising (such as "an older man is asked to resign his job to make room for a younger person"), we asked for a five-word answer to these two questions, "How did he feel" (*A*) and "What did he do" (*D*). Taking 20 of our collector hobbyist cases (our most and best known coöperative group), the projective test (which assumes that the subject identifies with each armature and so describes himself) was completed following all the other measures (*IVP*, questionnaire, and peer-count) previously described. Next an attempt was made to get a quantitative adjustment score for each subject in terms of his overall *A/D* quotient (ability to express *affect* over ability to get *defensive discharge* held between the bounds of rational control and complete inhibition). According to Sargent (5) the *A/D* might be regarded as "an index of psychic disturbance and mobilized discharge for warding it off." If this is true, then the optional mental adjustment relationship obtains when the *A/D* measure is in approximate balance "with a tendency for the best adapting subjects to respond to emotion-instigating aspects of a conflict situation with less need for defensive restitutional and controlling devices than the disturbed . . . at very high *A* levels, however, obsessive-compulsives and pre-schizophrenics occasionally give suspicion that *D* level controls are breaking even when the *A/D* remains high . . . in patient groups there is a consistent progression downward on the *A/D* quotient from the least to the most disturbed persons."

TABLE 5
MEDIAN CHI SQUARE TEST OF SIGNIFICANCE OF THE *IVP* VALUES TEST FOR 120 CASES
(A-C = 40 non-hobbyists — B-D = 80 hobbyists.)

	Amusement				Purpose		
M = .4	Gp. A-C	Gp. B-D		M = 4.8	Gp. A-C	Gp. B-D	
Above	(A) 24	(B) 40	NS	Above	(A) 12	(B) 50	p = < .01
Below	(C) 16	(D) 40		Below	(C) 28	(D) 30	
	Conformity				Non-Conformity		
M = 4.5	Gp. A-C	Gp. B-D		M = 3.5	Gp. A-C	Gp. B-D	
Above	(A) 14	(B) 36	NS	Above	(A) 27	(B) 32	p = < .01
Below	(C) 26	(D) 44		Below	(C) 13	(D) 48	
	Selfish				Social		
M = 5	Gp. A-C	Gp. B-D		M = 5.5	Gp. A-C	Gp. B-D	
Above	(A) 32	(B) 28	p = < .01	Above	(A) 14	(B) 40	NS
Below	(C) 8	(D) 52		Below	(C) 26	(D) 40	

TABLE 6
MEDIAN CHI SQUARE TESTS OF SIGNIFICANCE OF 120 CASES ON THE *IVP* INTEREST SCALES
(A-C = 40 non-hobbyists, B-D = 80 dedicated hobbyists.)

Scale A — Art-Music				
M = 7	Gp. A-C	Gp. B-D	Totals	
Above	(A) 11	(B) 62	73	p = < .01
Below	(C) 29	(D) 18	47	
Totals	40	80	120	
Scale B — Collecting-History				
M = 9.5	Gp. A-C	Gp. B-D	Totals	
Above	(A) 8	(B) 52	60	p = < .01
Below	(C) 32	(D) 28	60	
Scale D — Nature-Pets				
M = 10	Gp. A-C	Gp. B-D	Totals	
Above	(A) 20	(B) 38	58	NS
Below	(C) 20	(D) 42	62	
Totals	40	80	120	
Scale E — Mechanics-Science				
M = 9	Gp. A-C	Gp. B-D	Totals	
Above	(A) 20	(B) 34	54	NS
Below	(C) 20	(D) 46	66	
Totals	40	80	120	
Scale F — Sports				
M = 7	Gp. A-C	Gp. B-D	Totals	
Above	(A) 28	(B) 32	60	p = < .01
Below	(C) 12	(D) 48	60	
Totals	40	80	120	

TABLE 7

MEDIAN CHI SQUARE TEST OF SIGNIFICANCE FOR 120 CASES ON THE *IVP* PERSONALITY SCALE

(Gp. A-C = 40 non-hobbyists, Gp. B-D = 80 hobbyists.)

Energy-Activity Scale M = 5.5	Gp. A-C	Gp. B-D	
Above	(A) 4	(B) 56	p = < .01
Below	(C) 36	(D) 24	
Totals	40	80	
Emotionality Scale			
M = 7	Gp. A-C	Gp. B-D	
Above	(A) 24	(B) 44	NS
Below	(C) 16	(D) 36	
Totals	40	80	
Objectivity Scale			
M = 5	Gp. A-C	Gp. B-D	
Above	(A) 16	(B) 44	NS
Below	(C) 24	(D) 36	
Totals	40	80	
Sociability Scale			
M = 6.5	Gp. A-C	Gp. B-D	
Above	(A) 26	(B) 60	
Below	(C) 14	(D) 20	p = < .01
Ascendancy Scale			
M = 6	Gp. A-C	Gp. B-D	
Above	(A) 29	(B) 28	
Below	(C) 11	(D) 52	p = < .01
Totals	40	80	

TABLE 8

CORRELATIONS ON 20 COLLECTOR HOBBYISTS BETWEEN ADJUSTMENT (A/D) PROJECTIVE

Test Scores with Various Sub-scale *IVP* Test Scores	
(1a) with the dedicated interest sub-scale (collecting)	$r =$ .24
(1b) with all other interest sub-scales	$r =$ .08
(2a) with purposive life values	$r =$ .21
(2b) with conformity life values	$r =$ .10
(2c) with selfish life values	$r =$ —.18
(3a) with activity-energy personality factor	$r =$ .21
(3b) with emotionality personality factor	$r =$ .12
(3c) with objectivity personality factor	$r =$ .17
(3d) with sociability personality factor	$r =$ .23
(3e) with ascendancy personality factor	$r =$ .02

Since it is not our purpose to treat this present report from a clinical point of view, the main intent of the above quotation is to indicate the general rationale from which the writer "quantified" the *A-D* answers made by his 20 hobbyist cases to the five (conflict) armatures used. Number and degree on each scoring would still depend on the individual "set" of the scorer, hence at this development stage unworthy of further comment. Suffice it to say that of the four collector-hobbyists to whom the investigator gave the most unbalanced *A/D* scores (low affect and high defensive discharge) three were later found to have been under psychiatric treatment, and one (a true schizophrenic) was obviously so far out of line adjustmentwise that only the millions still at her command could be an excuse for bringing her out in public!

The primary matter which engages our attention here is not the degree of success or failure which pre-retirant hobbyists have met in stabilizing their ego-homeostasis in the process of aging. Rather we wish to see if any relationships whatever obtain between the crudely ascertained *A/D* scores of our 20 hobby group cases and more refined measures of the *IVP* test. The answer here is yes! Though not of any particular statistical significance, we show (Table 5) typical correlations found between *A/D* scores and (*a*) one interest sub-scale score (collecting) or all other interest sub-scale scores; (*b*) purposive, conformity, and selfish *Life Values Scales*; (*c*) energy, emotionality, objectivity, sociability, and ascendancy *Personality Scales.*

No attempt is made to draw any conclusion from the above. Results are included only to indicate that an important area in pre-retirement testing and counselling (i.e., relation of mental adjustment to interest and life values scales) yet remains an uncharted sea. The investigator has made preliminary excursions (using other trial-run adjustment measures than the *A/D* score) into this field and with even less promise of quantifiable results. Since funds for the present work are now exhausted, it seems most appropriate to conclude with a note on the need for further research.

H. SUMMARY CONCLUSIONS AND PROSPECT FOR FURTHER RESEARCH

In this report we have checked the validity of *IVP* test measures on a widely dispersed sample of pre-retirants belonging to one of several different avocational groups. We have found that scores on the interests portion of *IVP* are highly related to the field of avocational activity practiced by such pre-retirants. Furthermore, we have found indication that educational background and incidental training are significantly higher for specifically dedicated hobbyists than for those cases surveyed who were not dedicated to any particular avocational field. When we separate avocational cases on the basis of vocation practiced or trained for, we find an almost even split (50-50) between individuals whose vocation and avocation are very closely allied and individuals whose vocational vs. avocational interests bear little obvious connection. There is a slight tendency (well worth special study in later research) for pre-retirants of professional background (such as Doctors and

Scientists) to see the arbitrary retirement age of 65 as continued usefulness of a major interest in only slightly new dress. On the other hand, where the interest test pattern of such retirants is itself broad-based, through incidental training, many professional men and women are already developing a satisfying second vocation. Overall, we can say that the pre-retirant with a broad educational background and many interests appears to have more ready potential for finding his own healthful avocational retirement pattern than does a person of more limited background and restricted job experience. This does not mean, of course, that any illiterate bricklayer may not find his way into an unrelated pre-retirant hobby group (such as art or music) almost by accident and become very good at it. We simply do not yet know enough about how interests are created to be able to say, "Anyone can make a silk purse out of a sow's ear if he so desires." But we do suspect (on the basis of test, questionnaire, and peer-counts studied and herein reported) that aptitude and ability *are* important factors.

Turning now to the two other parts of *IVP* (Life Values and Personality Factors), we also found helpful suggestions for the pre-retirement counsellor. But test scores were here so often contaminated by interlocking attitudes and relationships that no simple adjustment formula emerges. In fact, only a skilled clinician with much experience in putting odd bits together intuitively could make much sense of a case which combined a high selfishness (Life Values) index and low ascendancy (Personality) index with an active do-good social-service interest; this avocation could be compensatory for either guilty vocational acts or for vocational inadequacies (at ease only in company that one can feel sorry for), etc. Such fascinating bypaths of inference are hardly germane to the present report. Several general differences in value and personality scores found between hobby and non-hobby groups do, however, require summarization.

1. On the average is there any evidence that the confirmed (and well-adjusted) hobbyist tends to move along a different life-value continuum than the non-hobbyist? Subject to the fact that the non-hobbyist group sample was very small, the answer is "yes!" Both *IVP* test scores and peer-count probing points to the average avocational centered pre-retirant as the more purposive (less amusement-bent in his life pattern) and still more the conformist than the non-conformist. Above all, he is less routinely selfish (but no more social-service conscious) than the average non-dedicated hobbyist.

2. On the average is there any evidence that differences in personality patterning separate dedicated hobbyists from the non-hobby group? With

much more question (partly because of recognized inadequacies of the personality test) here also we see a distinguishing trend. This presumes the average dedicated hobbyist to be more energetic activity-wise and more objective ascendance-wise than his counterpart. But the factor of emotionality or self-control (on which so much depends in good retirement adjustment) comes through in the present sample as a blank. There are just as many non-hobbyists as hobbyists who score high on this factor; in fact, the possibility exists that some dedicated hobbyists may be quite emotionally unstable even while "crazily good" in their hobby field. We must conclude therefore, that useful as *IVP* test leads may be, they do not give us any simple answer as to the quality or type of overall mental adjustment in the pre-retirant subject.

3. With further refinements can the *IVP* to be made to provide an overall adjustment index of aid to pre-retirement counsellors? The answer here must rest with further research. Basically (the writer believes), the problem is one of finding some objective screening test which will separate emotionally hazardous cases in obvious need of psychiatric or clinical attention from those of essentially sound emotional balance whose main retirement problem deals with interests and abilities, personal attitudes and the practical considerations of health, finance, or self-made opportunity germane thereto. In this report we have made mention of one limited attempt to develop such a screening device based on a projective Insight Test. Results there were encouraging though not significantly related to *IVP* test scores. Other existent screening tests of personality adjustment have also been tried but with even less available data even to hazard a guess as to their possible utility.

It is suggested that the hypothesis of *ego-homeostasis* advocated by Menninger, Sargent, and the writer now be made a basis for research on objective psychiatric screening devices tested along with *IVP* on large independent pre-retirant population samples regardless of educational background, hobby interest, or disinterest. Only thus might we find more adequate ways for moving the pre-retirant towards continued usefulness, greater happiness, and better mental adjustment in his years which still lie ahead. Such a program was not a part of the original grant for which this forms the terminal report. It can be only hoped that it may and will be carried on!

In conclusion, we may ask, "What utilization can one make of the *IVP* test as presently constructed?" One can, of course, turn to the clinician. However, the quantitative psychologist still believes that with properly constructed scales and additional evidence, the subjective character of much intuitive judgment can be reduced to a minimum. In this respect a trained counsellor would gain from our test many important leads both for discussing a man's retirement problems and for projecting a significant activity pattern into his later years. Our present experimental design may have overemphasized the importance of avocational guidance before it is too late; but we readily recognize that this plays only one part in the battle of ageing. To say that "if an oldster gets a hobby, everything else is taken care of" is sheer nonsense. Maturity, objectivity, and emotional balance are equally in the picture. "Grow up and you never grow old" said one wise counsellor before this "age of psychology" hit the testers. If, on the basis of test scores or otherwise, we psychologists can ever motivate pre-retirants to squeeze their present age completely dry of life-giving love, laughter, and achievement, they will be *glad* to move on to the normal next stage of social and emotional maturity. Forward looking enthusiasms for new interests, new friends, new service opportunities will then replace the backward look of futile regret which now discolors the mental adjustment of so many oldsters. And whether one retires (so-called) to things related with ones' past work or to something entirely fresh and different, thus will stand a true measure of the "age of psychology."[2]

REFERENCES

1. FREEMAN, G. L. A high level interest-values-preference test for counselling pre-retirants. *J. Gen. Psychol.* 1958, **46**, 121-139.
2. ———. The Energetics of Human Behavior. Ithaca: Cornell Univ. Press, 1948.
3. FREEMAN, G. E., & TAYLOR, E. K. The construction of a test battery for counselling mature individuals with psychological retirement problems. *J. Gen. Psychol.*, 1958, **45**, 133-151.
4. MENNINGER, K. A. The organism under stress-regulatory functions of the ego. *Amer. J. Psychoanal.*, 1953, **37**, 121.
5. SARGENT, H. The Insight Test. New York: Grune & Stratton, 1953.
6. SIEGEL, S. Non-Parametric Statistics for the Behavioral Sciences. New York: McGraw-Hill, 1956.

American Life Foundation
Watkins Glen, New York

---

[2] For further study of such points discussed on page 152, a useful recapitulation of Table 3's Chi Square Median Tests will be found in Tables 4-8. In these tables we have shown distributions made by hobbyists and non-hobbyists on the interests, values, and personality aspects of *IVP*.

$20.00 per annum
$10.00 per volume
$5.00 per half volume

QUARTERLY
Two volumes per year
Immediate publication

April, 1959
Volume 47, Second Half

(Founded in 1935 by Carl Murchison)

The Journal of PSYCHOLOGY

The General Field of Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold

under the restrictions of a liberal contract, and may not be resold for a period of five years from the date of publication.

APRIL, 1959

(Manuscripts are printed in the order of final acceptance)

Psychological aspects of confinement in fallout shelters 163
By J. Donaldson, K. Magnuson, L. McHugh, R. Niner, F. Watt, B. Williams, and A. Zink

Mental processes during the production of dreams 171
By Hugo G. Beigel

Phosphene threshold as related to age and sex 189
By Johs. Clausen and Rathe Karrer

First steps in the construction of a scale for the measurement of aggression 199
By Misha S. Zaks and Richard H. Walters

Validation studies of an aggression scale 209
By Richard H. Walters and Misha S. Zaks

The effect of varying intramuscular dosages of atropine and banthine on the galvanic skin response 219
By Daniel J. Perry, George E. Mount, and Chester D. Hull

(OVER)

Copyright, 1959, by The Journal Press
Provincetown, Massachusetts, U. S. A.
Entered as second-class matter January 26, 1937, at the post-office at Provincetown, Mass., under the Act of March 3, 1879

Two-year changes on the *ACE* by students in a college of art and architecture 223
By Frank J. Deignan

Differentiation and differential counsel of some neurotic personalities . . 231
By Edward S. Karon and Bertram P. Karon

The effects of flooding and physical suppression as extinction techniques on an anxiety motivated avoidance locomotor response 235
By A. Terrence Polin

An experimental study of the effect of head balance on patterns of posture and movement in man 247
By Frank Pierce Jones, Florence E. Gray, John A. Hanson, and D. N. O'Connell

Perception of verticality in middle and old age 259
By Peter E. Comalli, Jr., Seymour Wapner, and Heinz Werner

An investigation of the effects of reward and punishment on visual perception 267
By Frank S. Beatty, Lawrence E. Dameron, and Joel E. Greene

The behavioral world 277
By Harvey Nash

Comparisons of phenomenal distance in photographs of various sizes . . 289
By S. Howard Bartley and Harley J. Adair

MMPI screening of entering medical students 297
By Charles A. Knehr and Richard N. Kohl

Visual field and subjective center of a diamond 305
By Theodor M. Kunnapas

The influence of area on the critical flicker-fusion threshold 317
By William C. Roehrig

Ability vs. practice in paired-associate learning 331
By Clyde E. Noble

A further study in the consistency of judgments in categorizing verbal material 337
By M. E. Tresselt and M. S. Mayzner

Weber's law and related functions 343
By Gösta E. Ekman

Published as a separate and in *The Journal of Psychology,* 1959, **47**, 163-170.

PSYCHOLOGICAL ASPECTS OF CONFINEMENT IN FALLOUT SHELTERS*

Operations Research Office, The Johns Hopkins University

J. DONALDSON, K. MAGNUSON, L. MCHUGH, R. NINER, F. WATT, B. WILLIAMS, AND A. ZINK[1]

A. INTRODUCTION

No disaster in the history of mankind can parallel the devastation and turmoil that would result from a large-scale thermonuclear attack. The horror of such an attack would be compounded, beyond the intense initial shock, heat, and radiation, by long-term downwind fallout; "mysterious" radiation, which cannot be sense-percepted, would settle over hundreds of miles in lethal dosages. And, as her enemies' mastery of thermonuclear weapons and their means of delivery advances, the United States faces the possibility of such an attack.

Our active defense system is less than perfectly effective, i.e., at least part of an attacking force would accomplish its mission (2). Further, since much of our active defense efficacy depends on nuclear warheads, the very process of neutralizing the attack would result in fallout. Therefore, the Federal Administration, through the Office of Civil and Defense Mobilization, is currently promoting fallout shelters to protect the civilian population (3). (Fallout shelters, of value in the areas not devatated by the explosion, block fallout radiation. They are economically more attractive than blast shelters, which protect occupants from the initial shock and heat of the blast as well as all radiation.) A nationwide fallout-shelter system may become a reality. Should an attack occur, large groups of people might be confined for as much as two weeks or longer in these shelters under conditions radically different from anything in their experience (1).

The following discussion concerns some of the psychological problems which might arise in a large, public shelter occupied over an extended period of time by a heterogeneous group, although it is recognized that many of the same problems would exist in family shelters. Pre-attack means of easing some of

---

*Received in the Editorial Office on December 1, 1958, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This article is a summary of research performed by the authors while employed during the 1957 summer months at the Operations Research Office, Johns Hopkins University, Bethesda, Maryland. The study was originally published as Operations Research Office Staff Paper No. 45, March, 1958.

these emotional difficulties are proposed. The article considers an optimal situation, i.e., complete protection from fallout radiation, no physical damage to the shelter, and adequate material provisions. However, as an individual's physical condition affects his psychological state, a deficiency of necessities and the presence of radiation could only worsen the psychological problems of the shelterees.

By anticipating human behavior in a shelter situation following a thermonuclear attack and by considering this human factor in shelter-program planning, shelterees would more likely emerge in sufficiently good physical and mental condition to reconstruct the nation and continue the war effort.

Reaction to the attack and confinement in the shelter has, for this discussion, been roughy divided into two periods, traumatic and posttraumatic. The distinction between the two is not sharp; as the first phase (directly associated with the explosion) wanes, the second gradually assumes importance.

B. Traumatic Period

Immediately after a disaster occurs, victims tend to be dazed, with an absence of emotion or response. Although many who manage to reach a fallout shelter will not have witnessed extensive destruction, they will realize what is happening on the basis of previous knowledge.

At first the self-preservation instinct will be dominant; concern for personal survival may be the main stimulus for action until the people think they are safe. Everyone will probably be affected by this; however, for most people it should be of short duration.

As an individual's thoughts are directed away from himself, he will probably become intensely concerned with what has happened to his loved ones. Should the enemy attack at night, families will probably manage to reach the same shelter. However, should the attack occur during the day, family members will be scattered in several different areas, and each would be acutely anxious over the safety of the others. In this situation it is foreseeable that some people would make frantic attempts to locate relatives by trying to leave the shelter. Arguments or fights could arise over the question of exposing everyone to the contaminated air that would enter if the shelter door were opened. It would be necessary to convince the individual seeking to leave that little would be accompished and that he would probably die from radiation exposure or lack of food. A method of communication with other shelters would help alleviate some of this anxiety.

As attention continues to be diverted from concern for personal and family survival, shelterees will probably experience feelings observed in other disaster victims, i.e., self-reproach for not having done enough for those who perished. Other individuals may consider the disaster a punishment for their sins. Such feelings of guilt can lead to despondency and apathy, especially if shelterees have a great deal of time to brood over the catastrophe. Giving them time-consuming activities and the feeling that they are doing their part in shelter life would greatly relieve guilt complexes.

A reaction common to disaster victims is fear that the disaster will recur. Thus one of the most dominant fears of the shelter occupants could be that of reattack—another bomb or the invasion of enemy soldiers. Panic, a sudden and overpowering fright, would seem improbable, however, unless there is a terrifying incident such as a fire within the shelter. During the British air raids, the mass panic that authorities feared would result from herding people together in communal shelters never materialized.

Fallout shelters in or near the area directly affected by the blast will probably receive many injured surivors. Not only will wounded persons require medical aid, but they will have an adverse psychological effect on those who observe their injuries. It would be advisable to assign a section of the shelter to those with minor injuries, and imperative to have one for the seriously injured and the dead. Because some religious denominations are against cremation, preservation of dead bodies, perhaps by refrigeration, may be the only answer to the problem of their disposal.

Psychosomatic sickness can result from intense guilt feelings, unconscious needs for support and justification, or fear of physical illness. Fear of radiation exposure could cause the symptoms of a common shock case to be interpreted as radiation sickness. Some of the symptoms of radiation sickness, e.g., nausea, abdominal stress, vomiting, and diarrhea are also common to simple intestinal disturbances or fatigue. Certain odors, particularly of the dead, can also provoke some of the symptoms, as can reaction to danger. Pre- and post-attack education on radiation sickness would be the best combatant of this problem. Also, a means of detecting the amount of radiation inside and outside the shelter is imperative.

C. Posttraumatic Period

At first a large heterogeneous group in a shelter will be highly disorganized, the individuals in a state of dazed lethargy. But with the passage of time in the shelter, some of the problems directly associated with the thermonuclear attack, e.g., anxiety for one's life, will diminish in importance. When the shelter population begin to return to those habits connected with normal life, such as regular eating and sleeping, adaptation to their new environment should begin. With increased concentration on their surroundings, the shelterees will realize that their former way of life no longer exists. Homes, places of business and entertainment, and other possessions will have been destroyed; friends will have perished and social relationships will be disrupted; in fact, the United States government may be in complete chaos or may not even exist. Lacking orientation they may find no direction for their future life and goals and become, therefore, frustrated and apathetic. This point, by arbitrary designation, will mark the advent of the posttraumatic period.

It is during this period that the average individual will be most receptive to leadership. This can be of two types, either emergent or predetermined. Following catastrophic events a few persons with no previous leadership tendencies often become adept at organizing and managing people and situations—this is the emergent leader. However, emergent leadership is not entirely dependable. The first person arising to take control might be one with a formerly repressed desire for power. He might be thrilled with the idea of directing the masses, yet be incapable of leading well, lacking the necessary knowledge to make the proper decision. Furthermore, there is the possibility that such leadership will not emerge at all. It would probably be advantageous to predesignate and train key personnel, selecting people who have demonstrated leadership ability and who are respected by others in the community. However, these predetermined leaders might not reach their assigned shelters, or might prove incompetent under the strain of the attack. Should this happen, the shelter society would be dependent on emergent leadership.

The competency of the leaders and their effectiveness in organizing will determine how well the group as a whole accepts and adjusts to its altered conditions. Substituting group goals, problems, and activities for the shattered environment of the outside world is the challenge to the leaders. They must encourage a sense of reality and an interest in the future, drawing the shelterees out of their depression and apathy by making them feel important to the group, and by keeping them occupied so they cannot brood over the catastrophe. As many as possible should be given a job to do, no matter how insignificant. People who have occupations or avocations that can be quickly utilized, such as nurses, doctors, and clergymen, will find immediate places for themselves. Teenagers and elderly people can be given the job of caring for children; teachers can supervise recreational activities since they have had experience with children; individuals with hobbies might give instructive sessions to others. There are numerous other tasks to which the energies earlier expended in traumatic reaction to the attack can be diverted. Keeping the shelterees busy will also prevent boredom from creating major behavioral problems in the later periods of confinement.

Distrust of leaders and discord over rules could prove to be the problem of greatest consequence in the shelter. In the execution of certain rules, perhaps rationing or punitive measures, there may be question as to whether the leaders are acting in the best interests of the group. If the leaders receive communication from the outside, shelterees may suspect that information is being held back or distorted. Here, rumors caused by apprehensions and a lack of accurate information or distrust of the leaders, would prove very detrimental. To curb propagation of rumors, the leaders must gain the confidence of the people, and convince them that they are always told the whole truth; it would also be useful to have feedback, whereby various members of the group listen for and report any rumors, in order that they may be quickly resolved. Further, if partial self-government is permitted, the group will less likely resent restrictions, will feel an identity with the leadership, and will have greater trust in official reports.

Being considered a link with "official" authorities will lend considerable support to the leader's position; support from outside authorities would also combat illusions of isolation and abandonment. Conversely, if outside authorities exhibit incompetence, the shelter occupants may attribute this to the shelter leaders, representatives of the authorities. This would lead to ignoring of orders.

Whether conduct during shelter confinement will be good or bad will be largely determined by the shelterees' attitudes toward life. Some individuals' faith in God and/or humanity, and with it some of their ethical values, may be shattered by the disaster. If some of the shelter inhabitants adopt an "eat, drink, and be merry . . ." attitude, theft of food, disobedience of authorities, or sex problems may arise. If moral deterioration occurs, the consequences to the group as a unit will be severe. To prevent this deterioration, leaders should stress the idea that shelter life is a continuation of life on the outside, and that the shelterees are as much responsible for their behavior in the shelter as in their previous environment. With a large group concentrated in a limited area, social pressure could be a major influence to good behavior. Strong religious beliefs maintained throughout the disaster could also result in good conduct. Religious services would very likely prove one of the most beneficial shelter institutions, particularly if the shelter has been provided with devotional materials such as Bibles, crosses, and candles, which would increase the semblance of reality.

The desire for comfort and pleasure may not be especially important at first; when necessity demands it, people are usually willing to put up with privations. Just how long they will tolerate such inconveniences without being disturbed is not certain, however. Providing some small comforts such as candy and hot coffee might be worth the boost in morale. Decked bunks will do much to provide the privacy which some personality types may require. Additional privacy may be obtained by dividing large groups into smaller ones, perhaps by families or neighborhoods.

Should first aid and food be quite limited, however, animosities may arise over who is first in need. Mothers may say they are most in need on behalf of their children. Certain groups, such as the upper class, may claim that they are, by their nature, more deserving of aid than others. With very limited supplies, some may say that elderly people should not get the food or first aid since their life expectancy is quite short anyway. There is no absolute answer to this problem; shelter occupants will have to decide the question of sacrificing some members of the group so the rest may survive, on the basis of available supplies, forecasted time of occupancy, and ethical values.

During the initial periods of shelter life, anything so minor as personal irritations and social prejudices will be unimportant; people will realize that they have all experienced the disaster together, and a general feeling of altruism will probably prevail. As time passes during confinement, however, good feeling toward others may lessen and social problems may be magnified. Hostilities may arise along lines of former cultural or economic prejudices, and petty irritations may expand to the point of creating a discipline problem. It is likely that the group will choose scapegoats for venting irritation; complaining about inanimate items such as food would also provide the group with an emotional outlet. The leaders should emphasize the common situation and goals of the group, and de-emphasize its differences. Working with others is a strong unifying force.

As the shelter society becomes more stable, it will be harder for anyone to disrupt it. Stable conditions, however, are the culmination of the efforts of the leaders and may not be achieved. If the leadership proves ineffective during the early periods of shelter confinement, a disruption might very well occur, perhaps over the controversy of when to leave the shelter.

Pressure to abandon the shelter will increase with the passing of time. Because of ignorance of conditions outside, curiosity will be high. Concern for family well-being will further stimulate the desire to go out; remaining in a shelter to avoid the intangible danger of radiation may be an especially difficult precaution to take when highly concerned about loved ones. On the other hand, the shelter represents security and the comfort of society whereas the outside areas do not, being either devastated by the blast or contaminated by unseen fallout. Thus, many may want to remain. The authority of the leadership will again be tested on the question of whether conditions are safe for evacuation of the shelter. In this controversy, information from outside authorities on radiation conditions is mandatory. Also, if rubble from the blast prevents evacuation, shelter occupants may become highly disturbed unless rescue teams appear forthcoming.

Exit from the shelter will terminate the posttraumatic period, although the problems associated with it may continue for some individuals after they leave the shelter. Rescue or evacuation to other areas may at this time be possible; otherwise the shelter population may continue to inhabit the shelter, working in the area by day if radiation levels permit.

D. SUGGESTIONS

The shock of a thermonuclear attack would create numerous psychological problems; confinement in a shelter would intensify these and create new ones. Unchecked, these problems could result in "psychological casualties," individuals emotionally unfit to cope with the changed conditions and to rebuild the nation. Measures can be taken prior to any attack, however, which would help to alleviate these problems.

1. A means of communication with other shelters and outside authorities is imperative; it would aid in dispelling concern over family safety, preventing rumors, alleviating fears, solving problems in individual shelters, breaking down illusions of being the sole victims, and establishing a time for leaving the shelter. Since there is the possibility of enemy sabotage or control of communication lines, shelterees should not be completely dependent on this outside support, however.

2. By predesignating key personnel and developing standard operating procedures, there will be a better chance of efficient organization in the shelter. These operating procedures must be flexible enough to accommodate the variation in composition of the shelter group at different times of the day; they should further provide for incorporating emergent leaders and for shifting those leaders found most competent to positions of highest authority. In predesignating leaders, people should be chosen who hold positions of authority in normal life; shelter occupants will be more inclined to respect and obey them since they were formerly recognized as leaders. Even so, prospective leaders will require extensive civil defense training courses in the aftereffects of thermonuclear explosions, group dynamics, and first aid.

3. Doctors, nurses, psychologists, clergymen and others should be assigned to each shelter, if travel distances permit.

4. Americans must be thoroughly educated to the shelter standard operating procedures and to the effects of a thermonuclear attack, particularly radiation. They should be familiarized with the interiors of shelters by conducting air-raid drills, and by utilizing the shelters for other purposes, e.g., garage or recreation area. Knowledge is one of the most effective combatants of many of the psychological problems that would arise.

REFERENCES

1. BENTZ, R., *et al.* Some civil defense problems in the Nation's capital following widespread thermonuclear attack. *Oper. Res.*, 1957, **5**, 319-350.

2. COMMITTEE ON GOVERNMENT OPERATIONS. Civil Defense for National Survival: 24th Intermediate Report. Washington: U. S. Government Printing Office, 1956.

3. HOEGH, LEO A. Address before Dallas Federal Business Association, Dallas, Texas. Infor. Bull. No. 10, Fed. Civ. Def. Admin.; Battle Creek, Mich., June 18, 1958.

Operations Research Office
The Johns Hopkins University
6935 Arlington Road
Bethesda, Maryland

Published as a separate and in *The Journal of Psychology*, 1959, **47**, 171-187.

MENTAL PROCESSES DURING THE PRODUCTION OF DREAMS*

Department of Psychology, Long Island University

HUGO G. BEIGEL

A. PROBLEM

A number of studies made in the past proved that external stimuli affect dreams (1, 2, 3, 5). The extent of their influence, however, is not sufficiently known, nor has the process by which the sensory effect develops toward a story-like whole ever been observed. The present study attempts to fill these gaps.

B. METHOD

When a person sleeps in a moderately warm room and dreams that he is shivering from cold, the sensation and the subsequent imagery are sometimes explained as the result of a partial uncovering of the body. Probable though the causative relation is, certainty cannot be claimed unless the stimulus situation is definitely known and several dreamers experience similar effects under the same circumstances. These two conditions were met by exposing each of 16 volunteers to the same five stimuli, (*a*) five whistling sounds, (*b*) an open bottle of perfume, (*c*) turning on the ceiling light, (*d*) a cushion placed on the lap, (*e*) moving air from a fan. The sequence of the stimuli was varied and each stimulus was applied for 60 seconds. In addition, to make possible the observation not only of the crude effects of the stimuli but also their less obvious influences on subsequent mental processes, a common setting was established for every series of five stimulus-affected dreams. A natural dream reported by the subject in question was to serve as such. Against this mental background, it was hypothesized, any effect of the stimuli would stand out as a deviation. Each subject, therefore, was first asked to relate a recent dream; then he was hypnotized, the content of his dream was repeated to him and it was suggested that he would have the "same dream or a similar one." In every case, the dream was produced six times, the first time without application of a stimulus (so as to get a more complete picture of the original dream associations), the other five times while the afore-mentioned stimuli were introduced. After each dream the memory of the stimulus-influenced

---

*Received in the Editorial Office on December 1, 1958, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

version was erased and the original version was repeated to the subject. When the dreamer was finally awakened, he was informed about all the new features that had appeared and was asked to tell about any incidents, thoughts, or feelings which he considered related to them. Fifteen subjects reported while they were dreaming, one had to be awakened after every dream to make his report.

C. RESULTS

1. *Quantity of Responses*

All stimuli brought forth responses. Sixty-three dream versions reflected clearly the nature of the stimuli. Seventeen versions failed to do so. In six of these no dream was produced, in four the original dream remained unaffected, in seven the occurring changes could not be traced to the known sources and were therefore attributed to unknown variables. Fifteen dreamers reported reactions to the air from the fan, 13 to light, 13 to pressure. Twelve reacted with unmistakable responses to the whistle, 10 to the smell of perfume.

To measure the effect quantitatively, the output of sentences and phrases referring to new dream details was counted. Omitting the original 16 dreams, the 16 versions which were produced without stimulation, and in the remaining 80 versions all material that repeated features already mentioned in these two groups, 521 statements were recorded as indicating new dream experiences. Two-hundred and twenty-six of them referred to feelings, impressions, and ideas that manifested the effect of the stimulus situation as, for example, the response to light: "The sun is bright," or the response to moving air: "Wind and rain." Two hundred and ninety-five phrases which unquestionably depended on these responses but did not by themselves reflect such sensations were separated and will be discussed later. Reactions belonging to the first group varied in accordance with the stimuli applied. The number of pertinent verbalizations was distributed as follows. Positive responses to light 56 (25%), to wind 59 (26%), to pressure 56 (25%), to sound 37 (16%), to smell 18 (8%).

Although the number of verbal responses cannot reveal much about the dynamics of a dream, it gives some information about the impact of the effect of the stimulus. If a single response is given, as for instance, "The kettle is boiling" in reaction to the whistle, we know that the sensation has been perceived. When after this remark the subject reverts to the original dream without changing its features, the influence obviously was not great. When, on the other hand, the dreamer again and again mentions the felt sensation or elaborates on it, the result is unquestionably more persistent. An example is the dream with which Subject 5 responded to air from the fan.

> Swimming. I'm in the water. It was very hot. He dunked my head. Couldn't breathe. Dark. Couldn't come up. Cold. Wet. Very dark. Cannot breathe. He held my head.

Both the sensation "dark" and the feeling "wet" occurred rather frequently in response to moving air.

Beyond that, however, relatively little information can be obtained from a quantitative comparison alone. Not even the absence of a verbal response proves the ineffectiveness of the stimulus. Exposed to light, for instance, one subject writhed uncomfortably, but did not relate a dream. When she was awakened, she said that a sudden dazzling light had destroyed an incipient colorful imagery. In two other instances, the negative effect actually appeared to have special significance. Two male students, 20 years old, blocked when the cushion was placed on their laps although in every other situation they described their dreams fluently. The one, Subject 7, dreamed:

> Sitting at a beach. Watching the waves. The waves come up and recede. There is a fire on the beach. I sit near the fire. The waves hit the fire. The water comes up and goes back. . .

Then he hesitated and tried to arouse himself from sleep. The second, Subject 9, originally dreamed of being ordained for priesthood. He went through four dreams in a very lively manner, but when the cushion was placed on his abdomen, he fell silent. Asked whether he was dreaming something, he did not reply but woke up. Two possibilities for such zero-reactions obtrude themselves. One is that the stimulus aroused sexual ideas and that both men rejected incipient thoughts in that direction; the second—not irreconcilable with the first—is that the coming up and going back of the waves was in itself a response to the stimulus. Up-and-down movements were mentioned five times in this particular situation and never in any other. However, the sample is too small to present this reaction with surety as one typical of pressure on the lower body.

There is an additional point that relates to zero-responses. Some of the positive reactions were delayed to the last seconds of the allotted minute. This suggests that in some of the no-response cases traceable evidence of an effect might have come about if the period of stimulation had been extended. In other instances, an actual effect was not recognizable until the verbal reference to a locality mentioned in the dream was explained in the waking state. Thus Subject 4 reacted to wind: "Rye Beach. I'm in the Fun House."

Only in the subsequent interview was it learned that this remark alluded to an entertainment device in which the girls' skirts are suddenly blown up by blasts of air. Still other dreams took a course completely different from the original ones, but not until the end did it become evident that all the preceding imagery led up to the response that named the sensation and hence presupposed the influence of the stimulus. Subject 4, for instance, sleeping with a cushion on her lap, changed the dream locality from an apartment on the top floor of a house to a playground and said:

> It was really too low. On a seesaw. I'm walking on a seesaw. Bouncing. Down side goes up and other goes down. Till you fall. I fell. My girlfriend fell. She fell on a bar. Then another. And another. They are connected with . . . iron . . . to the wall. I fell against the bar with my stomach.

Falling with the stomach against a bar can be safely explained as a reaction to the pressure exerted on that body area. But if the dream had been interrupted before that information was given, it would have been difficult to recognize that the whole seesaw scene anticipated the effect of the stimulus.

These and other peculiarities encountered in the experiment afford us some insight into the mental machinery that tries to cope with an unexpected experience. Since they appeared to influence the formation of dream patterns in various ways, it is necessary to examine them in detail.

2. *Direct and Anticipatory Responses*

The response that gives the impression of greatest simplicity is what we may term the *direct response.* It appears so elemental because it coincides with the sensation known from the waking state. In reaction to light, the dreamer "sees" light or a glare or says, "Everything is lighter now." In response to the whistle, he hears whistling; in response to the stimulation from the fan he reports that it is windy or cold; in response to pressure he registers heaviness. However, to define such impressions precisely does not seem to be an easy task in a dream. Sensations were named directly 15 times, but only two times as first responses. In the other 13 cases, the verbalized classification was preceded by the sifting of apparently vague impressions. In response to whistling, Subject 4 dreamed:

> Come back from the beach. Went to my friend's house . . . motor . . . took the elevator to the top floor. Look out of the window. There is a playground. Looks like four buildings set in a square. I look down. Kids are playing. Noises. Whistles. Roller skates.

Like the seesaw and the falling of the girls in the previously cited dream, the playground, the kids playing, and the noises are *anticipatory responses.* The "motor" is in all probability an attempt to interpret the auditory impression. Only toward the end was the whistling picked up as a part of the sounds which the stimulus had conjured up.

The direct response, whether it occurred early or late, always named a sensation. Its perception remained either isolated or it generated additional ideas. These sometimes fell in with the preformulated pattern of the original dreams; sometimes they engendered a new pattern that was formed around the stimulus-promoted sensation. Subject 1 in her original dream was at "a dance around the school. Nobody dances but me. There are only two people. It was on the street. Seemed funny." Affected by the air from the fan, she dreamed: "I'm lying in the grass. At the side of. . . . The wind is blowing. I'm in the country. It tried to get dark. Then it was getting lighter again."

Anticipatory responses were always followed by others which more clearly defined the experienced sensation. Since these identifications inevitably followed the track indicated by anticipatory responses, the latter frequently set the scene of the new version. Even when the identification was wrong in terms of the waking experience and later was supplanted by the correct response, dreams sometimes slid back onto the course which the anticipatory response had taken. Out of the 26 primary responses, 45 belonged to this category.

3. *Interpretive Responses*

Anticipatory responses often gave the impression that in tossing up a tentative suggestion the cognitive powers of the brain were groping for the organization of unaccountable sensory phenomena into a *Gestalt* that made sense. A similar tendency became apparent in a slightly different form in another kind of dream responses, those that without naming the sensation related it to a specific source. The effect of the fan resulted in references to an open window or to "a blanket falling off;" the effect of light produced the reaction, "The sun was so strong." Static in the radio where no radio had been mentioned before explained the whistling sounds. In accord with their apparent function these responses were termed *interpretive.* Some of them were:

> Light—the sun is shining, daybreak, the glare from the snow, the bed lamp is on. Whistle—a train is coming, altar bells, they are singing, birds. Wind —Walking along the beach, rain beating down, sleeping on a roof in winter, getting wet, being dunked, flying in a plane, the summer ends, in a boat, on the ocean. Pressure—leaning against a rail,

> a squirrel is sitting on my lap, he put a blanket on me, the abdomen is bandaged. Smell—flowers, incense is burned, his hair has been treated with lacquer, etc.

In toto, 72 verbalizations indicated such interpretive attempts, answering, as it were, the questions What is happening? Whence comes this feeling? Why do I feel the way I feel? The answer was not necessarily confined to one interpretation. Subject 9, for instance, perceived his auditory impressions first as altar bells, then added, "They are singing the kyrie eleison." Sometimes a direct response followed the interpretation: "The sun is down. It is cooler." Sometimes the interpretation followed the naming of the sensation: "It's cool. A storm is coming up."

4. *Transposed Responses*

However, incompletely aroused from their sleep-sluggishness the cognitive powers often failed to interpret the sensory impressions adequately. Light brought forth the response hot; wind, with or without an intermediary reference to cold, the response dark. The cessation of light, on the other hand, produced not only darkness but also coolness. The response was transposed to another sense modality. Subject 14 reacted to pressure with kinesthetic sensations: "I can't walk. I drag myself across the floor." Other subjects reported more complex physical states such as dizziness, feeling sick and a tightening of the stomach in reaction to smell, and breathing difficulties in response to wind. Some of these transpositions occurred as immediate responses, some were elaborations following anticipatory or interpretive responses. Thus, discomfort because of perspiration appeared as an elaboration to "hot" (stimulus light), in another dream the sensation "hot" led to the feeling of thirst. Remarks referring to these two types of transpositions occurred 106 times. The distribution is given in Table 1.

5. *Elaborative Responses*

Transpositions, as the phrases exemplifying some of them made evident, named alleged sensations or interpreted them. Others belonged to a third group that was termed *elaborative responses.* These, in distinction to secondary elaborations, also reflected the stimulus situation but in a more remote manner. ("We took off our shirts" after the remark "It is hot"). At the same time, they enriched and expanded a dream situation that had already been established. The following example containing both elaborative responses and secondary elaborations was provided by Subject 3. His original dream ran: "I'm lying on the dock after fixing the motor of my boat . . .

Someone neatly dressed walked up behind me. A young girl. . ." Light changed several features.

> I'm sitting on the dock. It's hot. The sun is bright and very hot. I'm perspiring. Perspiring and dirty. There's a girl in a bathing suit. She looks fresh and clean. I want to go to the beach with her. Sun is bright on the eyes. I want to get out of the sun . . . etc.

The transposition of light to heat is followed by a state of physical discomfort as a result of perspiration. This feeling as well as the girl who now appeared in a bathing suit are considered elaborative responses. "She looks fresh and clean," is a secondary elaboration.

TABLE 1

Responses to:	Visual	Cutaneous	Kinesthetic and Balance	Auditory	Olfactory	Other body feelings
light	—	13	—	—	—	3
smell	4	—	2	—	—	4
pressure	9	—	8	—	—	10
sound	11	—	—	—	—	3
wind	22	—	4	4	—	9
total	46	13	14	4	0	29

Elaborative responses followed either correct or incorrect recognition of the pertinent sensation ("The sun was so strong—*we could not stand it without sunglasses. . . .*") Subject 6 reacted to pressure on the abdomen with the interpretation "a squirrel is sitting on my lap." Then she elaborated: "Who brought him in? I'm afraid. Push him away. I wouldn't push him." When after 60 seconds the cushion was removed, she continued, "Carol, thank you. She took him away. The squirrel. A squirrel scares me." Every sentence except "Who brought him in?" confirms the persistence of the feeling of pressure until the termination of the stimulation period brought relief and new adequate responses.

Elaborative responses, some of them at the same time transpositions, occurred 64 times. There were among them, however, 29 that merit separate mention. They were different from the transpositions previously discussed in that they referred to emotional states rather than to sensations or body feelings. Light produced the feeling of gaiety, hope, and elation; wind or the cessation of light that of gloom, depression, and hopelessness. The appropriate states were either named or implied. Subject 1, whose original dream depicted her lonely dance by the schoolhouse, responded to light:

> A dance. In the gym. A polka . . . Apple cider on the table. Noise. A lot of people. Sun all over. Everything is gay.

Some of the reactions appeared to be engendered by the symbolic connotations of the stimulus-produced sensations (light: "The sun is shining. There's a big meadow and some flowers. There's bluebirds singing. It's so relaxing and nothing unpleasant. I'm singing."); others were linked to body feelings. The discomfort resulting from the fan threw Subject 4 into a state of confusion. "Rye Beach. In the Fun House. Don't know which way to go. The boat is leaving. My friend. . . . Don't know how to get home. I'm looking for a way out." In particular pressure on the lap produced such reactions. The basic feeling apparently was that of restraint. From it, however, derived emotional reactions of various kinds. Subject 15 succumbed with resignation: "Running and running. And falling and falling. I keep thinking: I'm not going to get up again." Subject 3 in the same stimulus situation changed the resigned embarrassment of his original dream to anger.

> On dock. Getting nasty. Saw the girl with two other fellows. I got angry. Liked her very much. Wanted to hit one of them. Stood there watching. They are laughing. Wanted to hit one. She said hello to me. I was angry. . . .

Subject 13 in three dreams tried to find his wife. He suspected her of having deserted him. The version influenced by light turned his feelings: "Can't wait to go home. My wife has to be there when I get home." Subject 9 in his original dream was about to be ordained a priest. He described the line-up for the ceremony and his relatives in the church. Exposed to light, he said, ". . . on the pulpit. For the first time after many years of study. I read the gospel. It was a good sermon. The bishop smiled at me."

6. *Secondary Elaborations*

Secondary elaborations are, as some of our examples showed, all those pictorial, intellectual, emotional additions that round out the dream to a comprehensible illustration of a mood as do the birds and flowers in the cited dream of relaxation, the leaving boat that aggravated or justified the plight of Subject 4, the wish to go to the beach with the girl, and so on. Elaborations tend to carry the action forward; they insert contemplations or objections: "I want to get out of the sun. *I usually like it. . .*" Or in the squirrel scene: "Who brought him in?" Sometimes they try to explain associations that took an implausible turn. Subject 6 reacted to the fan with the following dream version:

> Bessie, would you please close the window. A knock on the kitchen door. Somebody. Get up and answer. You go. Put my coat on. Bob's standing there. You look like a wreck. He comes in. I told you I would

> come back. Are you here on leave? *A plane flew in for the weekend.* I'm happy to have you. A beautiful day. It's not that way in Iceland. All ice and snow. I was expecting a letter from you. I told you I'd come and get you. Who's there? *Bob just left for overseas.* Has to go out again. Come here. Let me brush my teeth and wash my face and wash my hair. I'm on his lap. Kissing him.

The attempts at logification do not always conform with what would be considered clear thinking in the waking state. Thus the delayed explanation "Bob just left for overseas" is undoubtedly somewhat confusing. But such reversals in the sequence of thoughts or events are typical of the thinking during the process of dreaming and must probably be attributed to the diminished efficiency of the brain's work during this state that is neither wakeness nor complete sleep. They contribute their share to the apparent irrationality of dreams, although at closer examination it can frequently be seen that few associations are entirely incoherent. They can usually be linked to a previous or a subsequent response or elaboration, just as the link from the stimulus wind to the suddenly appearing boyfriend in Iceland becomes plain only late in the dream when the conversation turns to the weather: "A beautiful day. It's not that way in Iceland. All ice and snow. . ."

The associative material presented in secondary elaborations, however, is subject not only to the push given by the stimulus-produced sensations but also to a pull that comes from other quarters. As a result it becomes instrumental in linking the configurations formed by interpretive, direct, and elaborative responses with a bigger pattern latently present in the dreamer's consciousness.

7. *Pattern Formation and Integration*

With the frame of reference provided for by the original dream, the forming of patterns in our experiment was not in all phases comparable to that of a natural dream. In six instances no pattern was formed, in seven cases the occurring changes did not seem to be attributable to the stimuli used, and four dreams ran their old course. In six dreams, on the other hand, multiple patterns developed, two of them engendered by the cessation of the stimuli, four of them striking a second theme that was unrelated to either the foregoing version or the original dream. Thus altogether 69 new patterns originated from the known outside force. In five of these, the effect registered but was absorbed without causing changes of importance. For example, Subject 14 exposed to the fan had the impression, "The window is open. It's windy," but then repeated her old dream.

In all other cases, the experienced sensations changed the dreams to various degrees. These changes were termed *minor* when the basic situation, action, and mood were preserved, but the sensation was so elaborated as to produce a new scene. Example: Subject 16 in her original dream visited a patient and gave him something to read. He praised her and remarked derogatorily on the books another nurse had brought him. Under the influence of light, the nurse followed the same action, but noticed, "The bed lamp is on. He isn't talking, only reading." Similar fleeting changes were recorded 10 times.

In other versions, one of the three criteria, the general situation or the action or the mood was altered, whereas the other two remained essentially the same. This was termed a *significant* change. Example: Subject 7 in his original dream was with a friend and a dog at Coney Island. They were standing on the boardwalk watching the ocean. Wind caused a change.

> Coney Island. A movie. In front of it a long line. It's going to rain. We are waiting in line to go in. It's raining. We are getting wet.

Significant changes occurred 19 times.

Still greater deviations from the original were termed *radical* changes. They were characterized by the alteration of two of the three criteria, action and mood, or action and general situation, or mood and general situation, while the third remained basically the same. Example: Subject 8 reported as his original dream:

> We were flying to another planet in a rocket ship. Two girls were with us. We were strapped down. Took off. There was a lot of pressure on us. Now we are on the real trip and land on the planet.

The version in response to wind ran:

> The sun is down. Evening. It's cool outside. I go and take a look. We must wait till morning. We must get used to the climate. Back to the ship. I take a thermometer reading. We go back to the ship.

Nothing but the allusion to the rocket ship was preserved. Radical changes were encountered 15 times.

Versions which were devoid of even one such reminder of the original dream were termed *total* changes. Twenty stimulus-produced versions proved to be completely unrelated to the original story. Example: Subject 5 reported as her original dream:

> There was a man. He was standing in a room. He had no face. It was dark. I had to go through the dark. He was standing in the light. I didn't know how to get there.

When a cushion was placed on her lap, she dreamed:

> I'm in the house of my sister-in-law. With my nephew. He is a little boy. There's a picture on the wall. He wanted to reach it. There's a lake on it, mountains, and snow. Come in. It's Barbie. She . . . I want to talk to him. About the boy. It's cold. Somebody put a blanket on me. I'm holding Johnnie.

Yet even when, as in this version, the original dream pattern was completely altered, material was sometimes found in the new dream that was related to the original even though it did not appear in it. There were on the one hand memories and happenings experienced when the dream had first appeared and, on the other hand, emotions which the dreamer considered causatively involved in it. The provenience of such material was ascertained in the interviews after the dreamers had been awakened. As a rule it was offered spontaneously when the subjects were confronted with new features in the versions they had produced. Thus Subject 2 related a recurrent childhood dream. In it he is standing in front of a dark alley. He must enter it because he has lost money that belongs to his mother. But he is afraid. "Because someone in there is sneaking around among the garbage cans. This someone is also looking for the money." To light the dreamer responded with a total change. "It's cold. Glaring snow. It's snowing. We broke the window. He is chasing me. He would kill us."

In the interview both elements mentioned became recognizable. Like the loss of the money the breaking of the window was a reality experience. On a sunny winterday the boys were throwing snowballs and broke the window of the laundryman's shop. The basic emotion in both dreams and reality experience was fear. The link from one dream to the other was established by the stimulus light, the effect of which was irreconcilable with the darkness of the alley and therefore shifted over the associations glare—snow to another fearful experience. Reality experiences relating to the original dream, though not apparent in it, were found seven times, emotional material to which the same conditions applied was expressed in 12 versions.

8. *Sources Outside the Original Dream*

This naturally raises the question: Whence came the material that could neither directly nor indirectly be traced to the preformulated dream pattern or to events or emotions related to them?

First consideration must necessarily go to the responses evoked by the selected stimuli. Clearly, once the stimulus provoked a change of mood or of locality, subsequent associations were forced to go on from there and sometimes they elaborated in a direction from which a return to the original pattern would have required a special effort. In some instances, consequently, the dream ended right there; in others it followed the path of associations that offered themselves more readily and thus was sidetracked toward ideas remote from the preformulated dream and its issues. The first-mentioned effect is exemplified by the brief lamentations in reaction to wind: "Sleeping on the roof. It's winter. The wind is blowing. The blanket is falling off." An example for the second was provided by Subject 8, who in his original dream had made the trip to another planet. To light he responded:

> We come out. The whole scene is lighter. It's summer. Hot. Take off our shirts. Start walking. Follow a road. Just walk. Two rocks. One on each side. It's dark in there. Walk in. Don't see anything. Out and back to ship. An apple tree. Wonder how it got here. Touch it. Feels different. Like cardboard. Knock on it. No sound. One girl wants to take an apple. Whole tree comes down.

In some versions, the memory of the original dream flashed up for a moment, but it appeared in an arbitrary and incoherent manner. For some reasons the path it indicated could not be followed. Thus Subject 2 in the dream where his fear related to breaking a window and being chased by the laundryman who "would kill us," stopped and after cessation of the stimulus added "He found the wallet."

The flash-back to a new facet of the original dream problem in a completely different set-up suggests a competition between two thoughts for dominance. A somewhat similar phenomenon was encountered when the stimulus-produced-sensation did not fit into the prevailing trend of thoughts. An adjustment had to be made. Three variations of such "adjustments" became apparent: (*a*) intense stimulation of that kind tended to inhibit thinking along any other lines except those relating to the stimulus; (*b*) the dream associations veered toward a more appropriate pattern that was equally well-rooted in the mind; (*c*) the experienced sensations were transferred to another person or another object. An example for this third type of managing the effects of the stimulus was given by Subject 16 (nurse visiting a patient). She worked the pressure on her abdomen into her dream in the following manner: The patient now was "sitting in a chair. He had an operation. He has bandages because of an appendix operation." Then she repeated the talk about the book and the colleague as before. Hearing whistling sounds in the next experiment not she but the patient became confused. "He doesn't know what is the matter. He becomes impatient. Tries to get the nurse's attention." Such transfers took place in eight instances.

Although among the changes that contained no material pertaining to the original dream several complete dream versions and many independent insertions had to be attributed in the first place to the influence of the stimuli, it was obvious that many of the elaborations were fed from other sources. The cardboard apple tree in the dream of Subject 8, for instance, may, like the darkness he sought between rocks, be an elaboration on light and heat (dry like coardboard) but it may be of a completely different origin. In some instances, the provenience of such remote material could be ascertained.

9. *Waking Experiences and Acute State of Mind*

An experience with Subject 1 was illuminating in this respect. The girl's original dream (dance at the schoolhouse) has been told before. When the whistle sounded, she was:

> standing at a station. Small. A train is coming. The train stopped. A man got off. He had a trench coat on. The train left. I watched. I was there all by myself.

The train in this changed version undoubtedly was associated to the whistle. A happening preceding the experiment helped to explain what gave this association priority over others. When the girl was asked to tell a recent dream, she produced the following: "There was a house and a man and I. There were also children and a train." Since this fragment appeared too scanty the subject was asked to think of another, more elaborate dream. She came up with the one that was then used. Apparently, however, the rejected material lingered in the girl's mind and associations to the whistle revived parts of it. It may, therefore, be justified to assume that in the selection of dream material any memories of thoughts or feelings may be given preference if they are close to the surface of consciousness because of their recency. The material may actually be insignificant but if it fits the sensations aroused by the stimulus or if it is usable as an explanation for these impressions, it is likely to be included in the dream.

Another factor in the utilization of otherwise unconnected ideas seems to be their proximity. Proximity may mean frequency of occurrence or intensity of experience or memory in the waking state. Frequency of thoughts and wishes has pobably—in this connection—the same effect as frequency of actual happenings. This hypothesis would explain the well-nigh stereotyped association of "beach," "water," and "wetness" to wind and the equally frequent association of "hot" to light, especially since the experiments were made in summer (though in a room the temperature of which was below 70 degrees).

Memories of previously experienced thoughts or happenings seemed to have a greater chance to be integrated with the newly forming dream pattern, however, when they were already organized into a tightly knit unit. The Fun House turned out to be a situation taken from reality; at the same time the girl remembered the shock she had experienced when the blasts of air blew up her skirt. A quarrel over a charge account on which Subject 6 dwelled at great length in one of her dreams also was a reality experience; her mother kept nagging her about it for two weeks, she explained.

Some of these reality experiences seemed to attain their priority because they were emotionally toned and the applied stimulus favored a response related to this particular emotional state. Subject 15, whose reaction to light elicited feelings of relaxation and hope, reacted to the whistle with visible discomfort and said:

> I'm in the city. Noise. Noise. I'm all alone. Nobody cares if I live or die. I don't care where they are. I think about my father. It's early in the morning. It's getting lighter and he isn't in yet. I wonder where he is.

The reality situation was this: The girl had been in New York for four months and could not get adjusted to the life there. The whistle roused associations of noise and these the latent feelings of discomfort and loneliness in the city. This again produced longing for home, where she was to go in a week. But in keeping with the original anxiety dream (running, falling, some one is chasing me, I don't know who they are) and in contrast with her earlier response to light, this expectation was clouded by the memory of another reality experience of her childhood when, filled with anxiety, she was waiting up until morning for father's return from one of his occasional drinking bouts.

Allusions to remembered reality experiences unrelated to the original dream were ascertained 12 times, allusions to acute emotional states, predominantly problems and conflicts, 16 times.

Naturally, the material illustrating such states or conflicts did not come exclusively from reality experiences. As being chased by unknown creatures is a symbolic presentation of anxiety or fear, so were the concerns in several of the experimental versions depicted by imaginary happenings. Subject 6, for instance, started with a simple dream story: "My girl friend Bessie is sleeping in my house. It was to late for her to go home by herself. She admires my new bed. I'm getting up to tell my mother that Bessie is here." We cited the experimental version in which her boyfriend appeared on the scene and assured her of his unshaken intention to "come and get" her. Although at the end of that dream she sat on his lap kissing him, it became clear that the deflection from the original had one of its causes in an emotional conflict which apparently was so important that it left its traces in three versions. In one, Bessie's boy friend showed up to take Bessie home ("I came to get you."). In the next version, this man was replaced by Bob, the dreamer's absent boy friend and the scene that we described took place. In the following dream (pressure) a quarrel developed between Bessie, our subject, and her sister. In its course, the subject was accused of being jealous. She replied: "I'm jealous? Because I have no boy friend?—I have." This indignant outcry was immediately followed by the scene with the squirrel, a change which contained not only unmistakable references to the stimulus but probably also symbolic allusions to sex. The impression that we were dealing with a sexual problem was further confirmed by a remark the subject made in a dream that belongs to a series of experiments concerned with the influence of body positions on dreams. There she said:

> I can't sleep. I have to show you Bob's picture.—Why not bring him around?—Don't tell. I'd rather he be with me than you.—Kathleen!—Well, you'd rather have Dick (the other girl's boy friend). Bob, in person, he doesn't stimulate me that way. . . .

Confronted with these fragments in the interview, the girl gave an account of her relationship to Bob. She had been upset by his impetuous demands and he had made no secret of his annoyance over her rigid resistance to his sexual advances. In fact, she did not expect him to return to her.

In all probability, the problem went even deeper than that, but it was not the experimenter's task to find out more about its nature. What made it interesting in the framework of our study was the fact that acute emotional material superseded the preformulated dream pattern and came to the fore in several versions which by themselves were different from each other in their main contents. Moreover, in every one of them the problem showed a different facet. The same phenomenon became apparent in the dreams of three other subjects. It suggests not only the conclusion that external stimuli may activate dreams of emotional significance but also another: that within the framework of the acute problem the aspect under which it is treated depends, to some extent at least, on the nature of the stimulation. According to the sensations or feelings which the stimulus evoked, hope or gloom, for instance, took the foreground. Subject 13 never repeated any feature of his original dream except its pessimistic mood. In two dreams, he took up this mood, whereas three showed no trace of it. In fact, light brought forth the anticipation of a happy ending. In two versions which Subject 15 produced the motif of the original dream showed up, but it turned, in accordance with the stimulus, once to hopeful thoughts (light: "everything will be all right in the future") and once (pressure) to complete helplessness and resignation. In a third version, the one quoted in the preceding paragraphs, two new problems came to the fore—her unhappiness in the city and the fear about father's state. It may be of special interest that in this version the original feature of being chased is alluded to but pushed aside as unimportant in the sentence "I don't care where they are."

The ultimate effect of the stimuli depends on several additional factors. There is, in all probability, the clarity or dimness of perception. It seems to make a great difference whether air can be discerned as wind and whether the origin of pressure can be plausibly explained to the self or whether they cause confusion. Eventually, of course, disposition and personality constellation play important rôles. They account for the repeated use of certain form elements such as presentations in dialogue or picture form, critical insertions and contemplation and possibly also for the frequent use of transfer of the experienced sensations to other persons. They are probably also responsible for the twice observed inclination to favor one mood and accordingly to turn the effect of any stimulus in its direction. These are issues still to be explored.

D. Summary

In observing the influence of five sensory stimuli on reported natural dreams it became possible to follow the mental processes that led from an initial groping of the mind for perception of more or less vague impressions to a meaningful organization of sensations, feelings, and thoughts and thence to the integration of these with a situational pattern containing also actions and emotions.

Applied to natural dreams, the results of our experiment indicate: (*a*) Stimuli perceived while a dream is in progress may change its mood or its content, depending on the feelings and memories associated with the stimulus-produced sensation. (*b*) They may inhibit the continuation of the dream. (*c*) External stimuli may actually start a dream. They do so by producing sensations which the brain powers are forced to organize into comprehensible *Gestalten* and subsequently into patterns that feed on memories of reality experiences, thoughts, past problems, and acute emotional states. By intensifying the reduced brain activity they rouse the sleeper to that state of half-awareness that is essential to bring about the phenomenon that is termed a dream.

References

1. CUBBERLY, A. J. The effect of tensions of body surface upon normal dreams. *Brit. J. Psychol.*, 1923, **13**, 243-265.
2. FREUD, S. The Interpretation of Dreams. New York: Macmillan, 1939.
3. MONROE, W. S. A study of taste dreams. *Am. J. Psychol.*, 1899, **10**, 326-327.
4. SELLING, L. S. Effect of conscious wish upon dream content. *J. Abn. & Soc. Psychol.*, 1932, **27**, 172-178.
5. SWEETLAND, A., & QUAY, H. An experimental investigation of hypnotic dreams. *J. Abn. & Soc. Psychol.*, 1952, **47**, 678-682.
6. WEED, S., & HALLAM, F. A study of dream consciousness. *Amer. J. Psychol.*, 1896, **7**, 405-411.

138 East 94th Street
New York 28, New York

Published as a separate and in *The Journal of Psychology*, 1959, **47**, 189-198.

PHOSPHENE THRESHOLD AS RELATED TO AGE AND SEX*

The Training School at Vineland, New Jersey

JOHS. CLAUSEN AND RATHE KARRER[1]

A. THE PROBLEM

When an alternating or intermittant current is applied across the eye through surface electrodes, the stimulation results in a sensation of flickering light. Such sensation, caused by non-photic stimulation, is referred to as phosphenes or more properly electrical phosphenes. Previous studies by the senior author as well as by other investigators (1, 2, 3, 6, 7, 8, 9, 10, 11) have led to the conclusion that the sensation is caused by direct stimulation of neural structures in the retina, such as the ganglion cells or the bipolar cells, rather than involving the photochemical system of the eye. It would appear then, that electrical phosphene thresholds indicate excitability of certain nerve fibers more directly than ordinary sensory thresholds where additional processes are integrated parts of the mechanism.

In previous studies (3) we found marked individual difference in magnitude of the phosphene threshold which may be interpreted as a function of individual difference in nerve excitability. In some cases age seemed to be related to phosphene threshold or excitability but other cases failed to confirm this. The number of cases included in these studies, however, was very limited. It appeared, therefore, reasonable to establish the relationship between phosphene threshold and age and sex, before venturing into a more elaborate study of determinants of the individual differences.

B. APPARATUS AND PROCEDURE

As a source of electrical stimulation a Hewlett-Packard Low Frequency Oscillator (Model 202 C) was employed. The sine wave signal was applied through two silver electrodes 25 mm. in diameter attached to the middle of the forehead and the right temple of the subject. The skin of these areas was cleaned with acetone and the electrodes covered with bentonite electrode paste and fastened with adhesive tape. The threshold was recorded in µA by means of a Hewlett-Packard Electronic Voltmeter (Model 400 B) with a 1000 ohm shunt resistor. Inserted in the circuit was a 10000 ohm resistor in series with the subject in order to reduce the effect of fluctuation in the subject's skin resistance. Also included was an interval timer which allowed a stimulus interval of exactly one second.

---

*Received in the Editorial Office on December 8, 1958, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This investigation was supported by Research Grant NSF-G 2106 from the National Science Foundation.

Previous studies have shown that the phosphene threshold is determined by current frequency and by the state of light adaptation of the eye (3). We selected, therefore, the frequencies of 20 and 50 cps. The first frequency was selected because it produces the lowest threshold. The 50 cps was included because it represents a fairly high frequency in the phosphene-producing band, while allowing threshold determination without appreciable discomfort to the subject. Two different states of light adaptation were attained by having the subject face a wide white screen, 7½ × 11 feet, illuminated so it reflected 0.03 and 30 mL respectively, as measured by a MacBeth Illuminometer. At each frequency both light conditions were used. Thus we recorded thresholds for four different conditions. All threshold determinations were obtained by using the psychophysical method of limits. The average of an ascending and a descending series constituted a threshold. For each condition the average of five thresholds was computed.

TABLE 1
AGE DISTRIBUTION OF SUBJECTS

Age	Males	Females	Total
10-14	13	13	26
15-19	6	10	16
20-24	10	14	24
25-29	11	10	21
30-34	6	1	7
35-39	4	2	6
40-44	6	6	12
45-49	2	6	8
50-54	3	6	9
55-59	3	1	4
60-64	3	2	5
65-69	1	3	4
Total	68	74	142

A total of 142 subjects were included in the study. Of these, 68 were males and 74 females. They ranged in age from 10 to 68 years. Table 1 shows the age distribution of the subjects. The subjects were all "untrained" subjects, meaning that with very few exceptions they were completely unfamiliar with phosphenes, and for many of them it was the first time they had taken part in a psychological experiment. There is reason to believe that the sample represents a fairly normal cross section of the population. Included were administrative and professional staff, hospital and laboratory personnel, teachers, playground supervisors, cottage parents, clerical staff, maintenance workers, etc., of The Training School. The younger subjects were recruited from the public schools.

In that this was the first time we had collected phosphene data on a large group of untrained subjects we were uncertain as to the reliability of the recorded thresholds. For this reason it was decided to determine thresholds on all four conditions on three separate days. After having obtained data from 31 subjects the data was analyzed and it was concluded that two sessions were sufficient (see below). Therefore, the rest of the subjects were tested in two sessions.

In the age and sex investigation we included four derived scores in addition to the four basic measures. The derived scores were: the ratios of the 20 to 50 cps threshold at .03 mL and at 30 mL; and the ratios of the .03 to 30 mL threshold at 20 and 50 cps. The first two ratios indicate threshold relationship to frequency and the last two ratios indicate threshold relationship to light adaptation. These ratios showed no relationship to age or sex and were therefore omitted from this presentation.

C. RESULTS

1. *Reliability*

As mentioned above, the first 31 subjects had thresholds determined on three different occasions for the purpose of determining reliability. The comparison of the three days for the four different measures are presented in Table 2. For one measure, the difference between the means of Days 2 and

TABLE 2
CORRELATION COEFFICIENTS AND CRITICAL RATIOS FOR REPEATED DETERMINATION OF PHOSPHENE THRESHOLDS

Comparison (N = 31)		.03 mL		30 mL	
		20 cps	50 cps	20 cps	50 cps
Days 1 and 2	r	.82	.88	.89	.86
	CR	1.07	3.14***	1.54	.77
Days 1 and 3	r	.86	.81	.81	.80
	CR	3.02***	1.13	.16	2.18*
Days 2 and 3	r	.86	.86	.93	.90
	CR	2.21*	.62	1.76	.73
(N = 114)					
Days 1 and 2	r	.72	.77	.57	.67
	CR	2.52**	1.52	.40	1.22

*Significant at .05 level of confidence.
**Significant at .02 level of confidence.
***Significant at .001 level of confidence.

3 is significant at the .05 level of confidence, but in view of the uniformly high positive correlations it was concluded that we could dispense with Day 3, and the remaining subjects were tested only twice.

For 114 subjects complete data were obtained on two different days. A comparison of the four thresholds from the two different days in terms of *r* and *CR* is presented in the lower part of Table 2. These correlation coefficients are definitely lower than those in the upper part of Table 2 but they are still highly significant correlations. One *CR*, namely 20 cps–.03 mL, attains significance at the .02 level, with the mean for the second day being lower than the first day's mean. This particular measure seems to be less stable than the remaining ones. The discrepancy in the correlation coefficients in the upper and lower part of Table 2 may be attributed to the fact that the group of 31 subjects have proportionally more subjects who have experience with experimental procedure and with introspection than is the case with the 114 subjects. The reliability in terms of *r* is of similar magnitude for males and females considered separately.

For our analysis of age and sex differences we included in addition to the above 114 subjects three subjects who for one or another reason did not have all four thresholds recorded on the first day, and 25 subjects on whom we were only able to get one set of recordings. We included this latter group in view of the high correlation coefficient between the two days, and the fact that for three of the four thresholds the difference between the first and second day did not attain statistical significance. Thus we have data from a total of 142 subjects which could be included for the purpose of ascertaining differences according to age and sex.

2. *Age Difference*

Plotting the four different measures for age groups in steps of five years gave a rather confused picture. Because of the small number of subjects in each group the ensuing curves had a rather irregular course. We therefore expanded the steps to 10 years, in an attempt to make the curve more regular. The result of this way of arranging the data is seen in Table 3. There is no obvious overall systematic change in threshold with age. It is interesting to note, however, that for both males and females there seems to be a gradual increase of threshold with increasing age for the three youngest age groups. Thus correlation coefficients were computed for the total groups and for the age groups above and below 35 years. The correlation coefficients for these groups are presented in Table 4. The majority of these coefficients are very low for the overall groups, for the age groups below 35 years and for the age groups above 35 years, on both males and females, separately as well as combined. There is a small positive correlation in the age groups below 35 which derives mainly from the male group.

TABLE 3
MEAN PHOSPHENE THRESHOLDS (IN μA) VERSUS AGE GROUPINGS

Age	N	.03 mL 20 cps	.03 mL 50 cps	30 mL 20 cps	30 mL 50 cps
Males					
Below 15	13	20.6	107.6	24.4	89.9
15-24	16	25.2	138.4	29.9	113.4
25-34	17	27.9	150.0	33.4	122.8
35-44	10	19.4	122.3	22.2	83.1
45-54	5	27.0	135.3	38.2	120.4
55-	7	15.3	127.9	22.2	88.6
Females					
Below 15	13	14.7	109.4	14.4	72.4
15-24	24	19.1	105.7	22.6	86.1
25-34	11	18.9	123.6	24.8	84.6
35-44	8	22.4	127.5	29.3	79.1
45-54	12	18.5	117.3	24.2	94.8
55-	6	22.9	144.9	20.2	100.7

TABLE 4
CORRELATIONS BETWEEN AGE AND PHOSPHENE THRESHOLD

		.03 mL 20 cps	.03 mL 50 cps	30 mL 20 cps	30 mL 50 cps
Males:	Total (N = 68)	—.11	.08	—.01	—.06
	Below 35 (N = 46)	.26	.36**	.17	.28**
	Above 35 (N = 22)	—.05	.10	.12	.19
Females:	Total (N = 74)	.14	.21	.08	.17
	Below 35 (N = 48)	.17	.12	.26	.09
	Above 35 (N = 26)	.00	.12	—.21	.20
Combined:	Total (N = 142)	.02	.14	.03	.03
	Below 35 (N = 94)	.24*	.27*	.23*	.22*
	Above 35 (N = 48)	—.01	.11	—.07	.20

*Significant at .05 level of confidence.
**Significant at .02 level of confidence.

Figure 1 is a typical example of threshold versus age which demonstrates the great threshold variability that is encountered in these determinations. Closer inspection of such data reveals that there are relatively few high threshold subjects which are skewing the means for each age group. If these extremely high threshold individuals are excluded, a difference in trend between males and females is accentuated. The mean thresholds of the males appears to increase to age 35 and then decline, while those of the females appear to increase to age 15 and then level off. If the population is divided into two groups of high threshold (above a value approximating the means) and low threshold subjects (below the same value), the trend is no longer present. The group of males and females do have in common a rise in threshold up to age 15. This difference in mean threshold of subjects below age 15 and those in the age group 15 to 25 is statistically insignificant for both males and females. The extreme variability in threshold which makes heavy demand on the sample size may be responsible for the trends in the data.

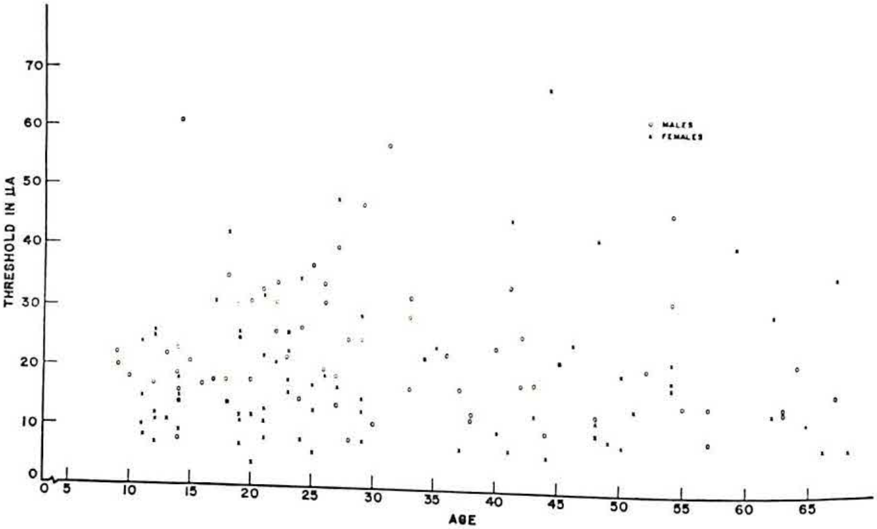


FIGURE 1
A plotting of phosphene threshold at 20 cps. .03 mL versus age.
A representative example of variation found in all four conditions.

3. *Sex Difference*

It appears from the data in Table 4 that there is a more pronounced sex difference in the younger age groups than in the older groups. Therefore group comparisons in terms of *CR* have been made for the total groups as well as for the groups above and below 35 years. The result of this comparison as well as means and standard deviations appear in Table 5. Table 5 shows that for the total groups as well as for the groups below 35, female thresholds are significantly lower than the male thresholds. A similar comparison for the groups above 35 shows no significant differences. The result for the total group is deceiving because of the influence of the younger subjects.

4. *Intercorrelation*

Independently for the two sex groups intercorrelations for the four thresholds were computed. These correlations are presented in Table 6. From this table it can be seen that all correlations are reasonably high, somewhat higher for females than for males. It seems reasonable to conclude that it is the

TABLE 5
COMPARISON BETWEEN MALES AND FEMALES: MEANS (IN μA), STANDARD DEVIATIONS AND CRITICAL RATIOS

	.03 mL		30.0 mL	
	20 cps	50 cps	20 cps	50 cps
		Total group		
Males N = 68				
M	23	132	28	103
σ	10.6	47.9	17.0	48.2
Females N = 74				
M	19	116	23	85
σ	11.6	44.8	15.1	38.8
CR	2.42	2.05	1.95	2.44
P Value	.001	.001	.001	.001
		Below 35		
Males N = 46				
M	25	134	30	110
σ	11.1	47.8	18.0	51.7
Females N = 48				
M	18	111	22	82
σ	9.3	43.1	14.0	39.5
CR	3.35	2.45	2.22	2.95
P Value	.001	.001	.001	.001
		Above 35		
Males N = 22				
M	20	127	26	89
σ	8.6	47.7	14.3	35.3
Females N = 26				
M	21	127	25	91
σ	14.7	45.9	16.9	36.7
CR	0.29*	0.0*	0.22*	0.19*

*Not significant.

TABLE 6
INTERCORRELATIONS OF PHOSPHENE THRESHOLDS IN MALES AND FEMALES (MALES IN ITALICS)

		.03 mL		30 mL	
		20 cps	50 cps	20 cps	50 cps
.03 mL	20 cps	——	*.60*	*.58*	*.52*
	50 cps	.67	——	*.52*	*.83*
30 mL	20 cps	.84	.62	——	*.59*
	50 cps	.55	.74	.60	——

same basic process which determines the thresholds, this being recorded for 20 or 50 cps, or for adaptation to .03 or 30 mL.

D. DISCUSSION

To obtain accurate threshold determinations it is necessary to employ as subjects people who have considerable experience in making the kind of judgment which is inherent in the task. The problem is particularly emphasized in our situation where the quality of the sensation represents a new experience to the subjects. It is, of course, impossible to find a sufficient number of trained subjects or to train a sufficient number of subjects, when one is concerned with the problem of the relationship of phosphene threshold to age and sex. It was, therefore, gratifying to find as good agreement between test and retest as Table 2 indicates. Our data, however, are not free from experimental errors. In that we recorded four thresholds which form a certain pattern, it was sometimes noted that a threshold would be definitely out of line. There is one type of deviation we encountered several times. At the first session thresholds for 20 cps at .03 mL and 30 mL could be readily obtained. In proceeding to 50 cps at .03 mL and 30 mL no threshold could be determined even at 300 μA. which was the highest stimulus intensity used. With two or three exceptions, all thresholds were obtained at the second session. We have no explanation for this initial failure to obtain threshold at 50 cps. Some of the subjects were, however, rather apprehensive about the test, since it involved what they thought was "electrical stimulation of the brain." At the second session they were considerably more at ease.

In spite of the above factors, we feel confident that we have obtained reliable phosphene thresholds on our sample of untrained subjects. Thus it seems that the main factor to be considered in evaluating our data is the extreme variability between individuals of the same age with its consequence for adequate sample size.

A change in light adaptation from .03 mL to 30 mL increases the 20 cps threshold, whereas it decreases the 50 cps threshold. The interaction between photic and electrical stimulation shows a different result at 20 cps than it does at 50 cps. This is true on the basis of the group means. If we consider the single individuals we find that 66.2% of the total sample has a higher threshold for 20 cps–30 mL than for 20 cps–.03 mL and 91.6% has a lower threshold for 50 cps–30 mL than for the 50 cps–.03 mL. The total pattern is found in 64.1% of the cases. This result obtained from such a large sample of untrained subjects confirms the result of a previous study utilizing a small sample of trained subjects (3).

Our data indicate a possible increase in threshold with increasing age to 15 and possibly 35 years. This tendency is indicated in the group means, more strongly for males than for females. Correlation coefficients in the male group support this tendency for all measures except 20 cps–30 mL, whereas the only correlation coefficient in the female group which supports this tendency is on the same 20 cps–30 mL measure.

When we consider the total age groups we see a rise and decline pattern for males and a short rise and levelling-off pattern for females. Although it is possible that there is an actual difference in pattern between the sexes, it seems at the present time more feasible to assume that they should follow the same pattern. It is particularly difficult to explain the rise and decline pattern in males in light of what is known about the effect of aging in other processes in man.

These findings, however, cannot be regarded as a final answer to the problem posed. As indicated in Figure 1, any age relationship is obscured by the marked interindividual differences. There are high threshold subjects and low threshold subjects in all age groups and the particular distribution of such people in each age group will determine the mean threshold of this group. The only way of circumventing this obstacle would be to have a tremendous number of subjects in each age group, or to do a longitudinal study. Neither of these possibilities is very feasible. The one thing that must be emphasized is that there are other processes, characteristic of each individual, which are much more important for the phosphene threshold than is age.

The significant differences between males and females were not anticipated. The result can hardly be fortuitous, such as would be caused by a chance occurrence of more low threshold females in our sample, because of the size of the sample in the group below the age of 35. The absence of a similar sex difference above the age 35 is puzzling indeed, and additional data should be collected before placing too much confidence in the results. We are at a loss for any explanation as to what in the female population would serve to produce lower thresholds.

Our data definitely focus interest on individual differences in the phosphene threshold. It has been demonstrated, by one of the authors, and others (3, 4, 5, 7, 11) that stimulus characteristics such as frequency, state of light adaptation, wave form, stimulus interval, etc., influence the threshold. We are now at the stage where we can change our orientation from the physical aspect of the stimulus situation to the physiological and psychological characteristics of the subjects which determine the phosphene threshold. We propose to study a group of subjects who are fairly well distributed over the range of threshold magnitudes in an attempt to find other measures which are related to phosphene thresholds. At present, we do not have any particular leads with regard to what these measures might be. It seems indicated that they might be looked for among those measures which are more closely related to sex than to age.

E. Summary

Phosphene threshold for four different conditions (two stimulus frequencies and two light conditions) were recorded for 142 subjects, ranging in age from 10 to 68 years. Repeated determinations showed high reliability of the recorded thresholds. Extreme variability of thresholds between individuals of the same age tends to obscure the relationship to age. Differences in relationship of threshold to age between males and females are indicated, but may be considered an artifact caused by the extreme variability. Females below the age of 35 seem to have lower thresholds than males of the same age. This sex difference does not seem to exist for the age groups above 35.

It is concluded that there are other factors, characteristic of each individual, which are much more important for the phosphene threshold than is age and sex.

References

1. Bouman, H. D. Experiments on the electrical excitability of the eye. *Arch. Nécrl. Physiol.*, 1935, **20**, 430-445.
2. ———. Electrical excitability of the eye: A study in clinical physiology. *Ophthalmologica*, 1940, **99**, 394-401.
3. Clausen, J. Visual Sensations (Phosphenes) Produced by *AC* Sine Wave Stimulation. Copenhagen: Munksgaard, 1955.
4. ———. Repetition rate and pulse duration as determinants of visual thresholds for electrical stimulation. *J. of Psychol.*, 1956, **42**, 329-339.
5. Gebhard, J. W. Thresholds of the human eye for electrical stimulation by different wave forms. *J. Exp. Psychol.*, 1952, **44**, 132-140.
6. Meyer-Schwickerath, G., & Magun, R. Ueber selektive elektrische Erregbarkeit verschiedener Netzhautantiele. *Arch. Ophthal.*, 1951, **151**, 693-700.
7. Motokawa, K. Visual function and the electrical excitability of the retina. *Tohoku J. Exp. Med.*, 1949, **51**, 145-153.
8. Motokawa, K., Iwama, K., & Endo, T. Ueber den Einfluss unterschwelliger elektrischer Reizung des Auges auf die Lichtempfindlichkeit der Netzhaut. *Tohoku J. Exp. Med.*, 1948, **49**, 331-338.
9. Rohracher, H. Ueber subjective Lichterscheinungen bei Reizung mit Wechselstroemen. *Z. Sinnesphysiol.*, 1935, **66**, 164-181.
10. Schaefer, H. Elektrophysiologie: II. Wien: Deuticke, 1942.
11. Schwarz, F. Quantitative Untersuchungen ueber die optische Wirkung sinusfoermiger Wechselstroeme. *Z. Sinnesphysiol.*, 1940, **69**, 1-26.

Training School at Vineland, New Jersey
Landis Avenue & Main Road
Vineland, New Jersey

Published as a separate and in *The Journal of Psychology*, 1959, **47**, 199-208.

FIRST STEPS IN THE CONSTRUCTION OF A SCALE FOR THE MEASUREMENT OF AGGRESSION*

Department of Neurology and Psychiatry, Northwestern University Medical School, and Department of Psychology, University of Toronto

---

MISHA S. ZAKS AND RICHARD H. WALTERS[1]

A. INTRODUCTION

There is little, if any, disagreement among students of human behavior that aggression signifies an important aspect of human conduct. Various aspects of aggression such as its function as a personality variable and its relationships to mental disorders, interpersonal relations, and frustration tolerance have been stressed by numerous psychologists (1, 9, 12, 13, 14, 15, 19, 22, 23). The recent literature has been reviewed by L. Berkowitz (2).

It is obviously important to evaluate the extent of aggressive tendencies in order to be able to predict and to control their overt behavioral manifestations. The clinical psychologist often attempts to determine such tendencies from projective test data. His estimates are usually based on vague clues in the Rorschach and *TAT*. Validity in such cases depends entirely on the individual clinician's ability and experience or on the so-called "clinical intuition" of the interpreter. Various formal devices for assessing aggression from projective tests have been devised and reported (7, 10, 11, 16, 17, 20, 21). Attempts to predict overt behavior by means of these devices have yielded inconclusive results. While indications of aggression on the projective test level may provide some useful clues for clinical diagnostic purposes, such estimates too often fall short of the necessary standards for research in the behavioral sciences.

---

*Received in the Editorial Office on December 11, 1958, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The authors wish to express their appreciation for the assistance and coöperation of Cook County Sheriff Joseph D. Lohman, Cook County Jail Warden Jack Johnson, Assistant Warden Hans Mattick, and members of their staff; the Superintendent and staff of the Training School for Boys, Bowmanville, Ontario; the Director of Education and members of the staff of the Scarborough, Ontario, Board of Education; Canada Packers, Ltd.; the following members of Northwestern University School of Medicine, Department of Neurology and Psychiatry, Chicago, Illinois: Benjamin Boshes, M.D., Professor and Chairman; George K. Yacorzynski, Ph.D., Professor and Chief, Division of Psychology, and Jerome Cohen, Ph.D., Associate Professor and Research Coördinator, Division of Psychology.

The authors are indebted to Dr. Leo Srole for permission to use, in this study, his scales, some of which have not yet been published.

In order to facilitate research on aggression objective instruments are needed. The advantage of such instruments is that their claims to validity and reliability can be readily determined, and the conditions under which such tests do or do not discriminate can be delineated, and subsequently confirmed or refuted through the replication of studies. The need for instruments designed specifically for the measurement of aggression has been recognized by others. In addition to the Rosenzweig *P-F* Study (23) and older inventory-type techniques, there are at least three recently developed instruments (4, 8, 24) that consist of items selected from the *MMPI*. In these inventories the items used as indicators of aggression are selected on the basis of the agreement of judgments made by clinical psychologists. This method of selection was used by Taylor (26) in the development of her Manifest Anxiety Scale (*MAS*). Such a method does not, however, by itself guarantee validity; moreover, its automatic application to the measurement of aggression is open to criticism.

A large proportion of items selected by clinicians as indicating aggression reflect the thought content of paranoid patients. While paranoia undoubtedly has a strong hostility component, the most frequently encountered hospital patients with pronounced paranoid ideation are schizophrenic, and the distortions in their attitudes are part of a pervasive affective and ideational disorganization. Although items capable of detecting pathological aggressive responses similar to those found in paranoid patients are undoubtedly applicable to a small fraction of non-schizophrenic subjects, their value in an aggression scale for wide and general application is questionable.

Taylor (26) made use of college students for the development and validation of the *MAS*. The disadvantages of her procedure are evident from subsequent studies. When Hoyt and Magoon (18) attempted to validate the *MAS* for a college counseling population, only 30 of 50 *MAS* items seemed to be related to anxiety. Moreover, in a cross-validation study on neuropsychiatric patients (3), the number of discriminative items shrunk to 14, i.e., only 22 per cent of the original 65 *MAS* items had discriminative value when cross-validated on subjects other than college students. There is no good *a priori* reason to believe that, of aggression items selected in the same way as *MAS* items, more than one in five would possess general validity. In any case, clinical agreement cannot in itself be considered an impressive criterion of validity for any type of item.

Buss and Durkee (4) have proposed an inventory capable of assessing hostility of different kinds. A factor analysis revealed two factors: "an attitudinal component of hostility," consisting of resentment and suspicion, and a "motor" component, which consisted of assault, indirect hostility, irritability, and verbal hostility. Buss and Durkee emphasized that while the (college student) population used in deriving the two factors was normal, the factors also appeared relevant for the study of clinical populations, as the factor comprising resentment and suspicion items seemed to characterize ideation frequently expressed in paranoid conditions. Their results suggest that indirect hostility, irritability, and verbal hostility are sufficiently intercorrelated to form a single psychological factor of aggression characterized by its "motor" component, i.e., a tendency toward motor or behavioral manifestations. Items measuring aggression of this kind should not be confused with items of a quite different type (frequent in clinicians' judgments of hostility), which are heavily weighted with paranoid ideation.

B. PROCEDURE

The aggression scale described in this report was developed from a selection of items which the senior author believed to be related to aggression. Some items were found in existing inventories, where they were utilized for various purposes; they were, if necessary, modified for the purposes of this study. Other items were constructed on the basis of the author's speculations about the nature of aggression.

In general, items included were of three types: (*a*) those dealing with some aspect of aggression, as determined by manifest content, e.g., "I often feel like picking a fight with someone"; (*b*) items related to aggression in an indirect manner, e.g., "Most people get killed in accidents because of their own reckless driving"; and (*c*) items for which agreement by the testee was assumed to denote aggression on theoretical grounds rather than on account of their content, e.g., "I almost never dare to express anger towards people for fear I may lose their love or approval." In all, 33 such aggression items were interspersed with other items in a 160-item inventory developed in connection with an ongoing study of emotional problems in heart disease.

The inventory was administered under the title, *The Personal Opinion Inventory* (POI). It consists of two parts. Part One contains 70 items which are of the indirect attitude measurement type described by Campbell (6). Included in this part are an abbreviated version of the California *F*-Scale (11 items), and three scales devised by Srole (25)—the Anomia Scale (10 items), the Social Prejudice Scale (7 items), and the Rigidity Scale (6 items). These scales are scattered and interspersed with 36 buffer items. The buffer items were especially prepared for the indirect estimation of aggression and other behavior tendencies. Part Two of the inventory consists of 90 items of the direct or self-descriptive type. Included in this part are 18 *MMPI* depression items, and the 14 *MAS* items validated by Buss. Other items were adapted or newly devised, and are currently being studied in order to assess suicidal tendencies, dependency, sexual inhibition, etc.

In order to establish the discriminative possibilities of each of the 33 tentative aggression items, a group of subjects with known high aggressive tendencies was chosen. This group consisted of 22 inmates of a midwestern prison. Each subject was given a battery of tests including the *POI*, projective drawings, the Rorschach, and some *WAIS* scales.

These prisoners were repeated offenders with at least two convictions for such crimes as assault and battery, assault with intent to kill, assaultive robbery, etc. Since 19 of the 22 prisoners were Negroes, it was decided to include only the Negro subjects in the criterion sample. The answers of the assaultive group to the 33 aggression items were compared with those of a preliminary control group consisting of 60 white adult middle-class subjects. Nineteen items discriminated at the .01 level. Only these items were selected for further study.

The assaultive group was now compared with 100 Negro adult subjects residing in Negro areas of Chicago; this latter group, like the assaultives, was almost entirely lower-class in socio-economic status. A second control sample consisted of 100 white adult males; included in this group were 75 males enrolled in a trade school in a Negro-white fringe area of Chicago and 25 adult evening students attending a downtown *YMCA* high school. At the time of testing, most control subjects were unskilled or semi-skilled workers.

Of the 19 items that discriminated assaultives from the middle-class sample, 7 items showed no discriminative value when the control groups of much more comparable socioeconomic and educational levels were employed. These items were consequently dropped from the scale.

Table 1 lists the 12 items that were retained and the proportion of each group that agreed with each item. On the left is the number of each item as it appears in the 160-item inventory.

All 12 items discriminated at a better than .05 level between the assaultive group and the combined normal control group. Ten items discriminated both between the assaultives and the Negro controls and between the assaultives and the white controls. Item 24 discriminated both between the assaultives and the Negro controls and between the assaultives and the combined control group; for the white sample the difference was in the predicted direction,

TABLE 1

PROPORTION OF ASSAULTIVES AND NORMAL CONTROLS AGREEING WITH AGGRESSION SCALE ITEMS*

Item No.	*Statement*	Negro assaultives N = 19	Negro controls N = 100	White controls N = 100	Combined controls N = 200
5	There are two kinds of people in this world: the weak and the strong.	95	67	43	55.0
22	Dealings with policemen and government officials are always unpleasant.	53	25	20	22.5
24	Most people get killed in accidents because of their own reckless driving.	79	53	68	60.5
27	Horses that don't pull should be beaten or kicked.	26	06	07	6.5
40	At times we enjoy being hurt by those we love.	58	26	22	24.0
44	Many a decent fellow becomes a crook or a criminal because he can't stand to be pushed around so much.	84	66	67	66.5
67	I easily lose patience with people.	53	24	29	26.5
71	I often do things which I regret afterwards.	73	40	54	47.0
93	It makes me mad when I can't do things for myself the way I like to.	95	54	64	59.0
101	Occasionally I was in trouble with the police or law.	79	17	15	16.0
108	I almost never dare to express anger toward people for fear I may lose their love or approval.	58	35	25	30.0
110	As an adolescent (or young kid) I often mixed with the wrong crowd.	68	38	26	32.0

*Agreement with any item is given a score of one, disagreement a score of zero.

but failed to reach significance. Item 44 fell just short of discriminating between the assaultives and each of the control samples taken separately, but discriminated when the control samples were combined.

In subsequent studies the 12 aggression items were treated as a provisional scale. In order to determine the validity of this scale within the normal range of the population, two experiments were devised. These experiments are reported in another paper (28).

An estimate of reliability of the aggression scale was obtained by administering the *POI* to 20 undergraduate students. Three weeks later the inventory was administered to the same students. The inventory was first given as a group test, the instructor reading off the items at 15-second intervals. The second time it was administered as an individual test. A product-moment correlation coefficient of .86 was obtained for the two sets of scores, indicating a satisfactory degree of reliability. It is evident from this test-retest correlation that the scores are not significantly influenced by the form of presentation of the inventory.

C. FURTHER INVESTIGATIONS

It seemed reasonable to assume on grounds of clinical experience that adolescents tend to manifest more aggression than adults. To determine whether the provisional aggression scale would reflect a weaker inhibition of aggressive tendencies in adolescents, 72 Toronto high school boys between the ages of 14.0 and 16.11 were compared with 37 adult male employees, aged 30.0 through 54.11, of a Toronto industrial concern. As Table 2 indicates, there was a highly significant difference between the adolescent boys and the adult males.

TABLE 2
COMPARISON OF GROUP MEAN SCORE DIFFERENCES BETWEEN ADOLESCENT BOYS AND ADULT MALES ON THE AGGRESSION SCALE

	Age range	*N*	*M*	*S*	*t*	*p*
Adolescents	14.0–16.11	72	6.15	1.91	12.92	< .001
Adults	30.0–54.11	37	2.11	1.11		

One might assume that an instrument that measures aggression for the adult population should also reveal aggressive tendencies among adolescent boys. Since the *POI* aggression scale was developed on adult samples only, a comparison of aggressive adolescents with normal adolescent controls could provide additional validating evidence. The inventory was consequently administered to 72 delinquents, aged 14 through 16, from a juvenile training school in Ontario. Boys are not ordinarily committed to this training school by the juvenile court unless they are repeated offenders. The aggression scores of the 72 institutionalized delinquent boys were compared with scores obtained from a group of 72 Toronto high school pupils, matched with the delinquents for age. Table 3 indicates that the scale discriminated significantly between these two groups.

TABLE 3
GROUP MEAN DIFFERENCES BETWEEN INSTITUTIONALIZED JUVENILE OFFENDERS AND HIGH SCHOOL CONTROLS

	N	*M*	*S*	*t*	*p*
Institutionalized offenders	72	7.61	1.54	5.01	< .001
High School controls	72	6.15	1.91		

Since the control subjects were taken from an industrial development area in which there is a relatively high incidence of juvenile delinquency, it was not surprising to find that 54 per cent of the high school control group answered "Agree" to Item 101 ("Occasionally I was in trouble with the police or law"). This item is perhaps sufficiently ambiguous to reflect differences in interpretation within non-institutionalized samples. However, if this item is given a legal and strictly literal interpretation, one can describe the 39 high school boys who agreed with the item as "self-confessed delinquents." While these high school students admit occasional troubles with the police or law, it is reasonable to asume that most of them do not act out aggressive impulses in the extreme forms found among institutionalized juvenile offenders. Therefore, the scale should still discriminate when institutionalized offenders are compared with "self-confessed delinquents" who manage to adapt to the demands of a regular high school. As Table 4 shows, this expectation was fulfilled (Item 101 held constant).

TABLE 4
GROUP MEAN DIFFERENCES BETWEEN INSTITUTIONALIZED JUVENILE OFFENDERS AND "SELF-CONFESSED DELINQUENT" BOYS ATTENDING REGULAR HIGH SCHOOL

	N	*M*	*S*	*t*	*p*
Institutionalized offenders	72	7.61	1.54	2.45	< .01*
Self-confessed delinquent High School boys	39	6.81	1.58		

*One-tailed test.

If it is assumed that the various kinds of antisocial activities have a hostility component, the applicability of the aggression scale to at least one other form of asocial behavior has already been demonstrated in a recent study of heroin addiction. Table 5 gives the mean aggression scale scores of groups of Negro addicts actively indulging in heroin intake, jailed Negro addicts (forcibly withdrawn), Negro non-addicts, and white non-addicts. The groups were matched for age and educational level.

Negro heroin addicts, whether tested in their ecological areas or under

TABLE 5
MEANS AND STANDARD DEVIATIONS OF ACTIVE NARCOTIC ADDICTS AND JAILED NARCOTIC ADDICTS WITH NEGRO AND WHITE CONTROLS ON THE AGGRESSION SCALE

Measures	Active addicts N = 35	Jailed addicts N = 35	Negro controls N = 35	White controls N = 35
Mean	6.00	6.26	4.94	5.00
Standard deviation	1.76	2.02	2.29	1.91

forced withdrawal conditions in jail, gained significantly higher scores than the matched Negro and white control groups (Table 6).

TABLE 6
GROUP MEAN COMPARISON BETWEEN ACTIVE *N-A*'s, JAILED *N-A*'s, AND NEGRO AND WHITE CONTROLS ON THE AGGRESSION SCALE

Group compared	t	p
Active *NA*'s vs. Jailed *NA*'s	0.57	N. S.
Active *NA*'s vs. Negro controls	2.14	$< .05$
Active *NA*'s vs. White controls	2.27	$< .05$
Jailed *NA*'s vs. Negro controls	2.52	$< .05$
Jailed *NA*'s vs. White controls	2.62	$< .01$

The fact that the aggression scores remain elevated in the addict groups after withdrawal may indicate that aggression is a relatively stable characteristic of lower-class Negro addicts.

D. SUMMARY AND DISCUSSION

A 12-item aggression scale was developed from 33 items interspersed among other items in a 160-item inventory. The general procedure consisted in comparing the responses of a group of assaultive criminals with a series of control groups and then proceeding to further comparisons that might indicate that the scale was of more general applicability. The preliminary evidence suggests that the scale fulfills the expected requirements of reliability and validity.

Of the 33 items originally selected for study, none of the "obvious" self-descriptive items dealing with aggressive content attained statistical validity. Such items as "I often feel like smashing things," "I often feel like swearing," "I often feel like picking a fight with someone," and "Somehow, I often seem to hurt persons I love," did not discriminate between supposedly aggressive and supposedly "normal" groups. Nor did self-descriptive items about such themes as liking guns, liking to tease friends or animals, liking to watch fights or wrestling matches, or to read reports about crimes. No items denoting participation in, or identification with, some directly aggressive activity, physical or verbal, discriminated in this study.

The available evidence indicates that the 12-item aggression scale is applicable to the study of aggression within the general population as well as to its antisocial manifestations in groups charactedized by assaultiveness, narcotic addiction, or juvenile delinquency.

REFERENCES

1. ADORNO, T. W., FRENKEL-BRUNSWIK, E., LEVINSON, O. T., & SANFORD, R. N. The Authoritarian Personality. New York: Harper, 1950.
2. BERKOWITZ, L. The expression and reduction of hostility. *Psychol. Bull.,* 1958, **55**, 257-283.
3. BUSS, A. H. A follow-up item analysis of the Taylor anxiety scale. *J. Clin. Psychol.,* 1955, **11**, 409-410.
4. BUSS, A. H., & DURKEE, A. An inventory for assessing different kinds of hostility. *J. Consult. Psychol.,* 1957, **21**, 343-349.
5. BUSS, A. H., DURKEE, A., & BAER, M. B. The measurement of hostility in clinical situations. *J. Abn. & Soc. Psychol.,* 1956, **52**, 84-86.
6. CAMPBELL, D. T. The indirect assessment of social attitudes. *Psychol. Bull.,* 1950, **47**, 15-38.
7. CHILD, I. L., FRANK, K. F., & STERN, T. Self-ratings and *TAT*: Their relations to each other and to childhood background. *J. Pers.,* 1956, **25**, 98-114.
8. COOK, W. W., & MEDLEY, D. M. Proposed hostility and pharisaic-virtue scales for the *MMPI*. *J. Appl. Psychol.,* 1954, **39**, 414-418.
9. DOLLARD, J., MILLER, N. E., DOOB, L. W., SEARS, R. R., FORD, C. S., HOWLAND, C. I., & SOLLENBERGER, R. T. Frustration and Aggression. New Haven: Yale Univ. Press, 1939.
10. ELIZUR, A. Content analysis of the Rorschach with regard to anxiety and hostility. *J. Proj. Tech.,* 1949, **13**, 247-284.
11. FINNEY, B. C. Instructions for the Use of the Palo Alto Aggressive Content Scale. Palo Alto, Calif.: Veterans Administration Hospital, 1951 (mimeographed).
12. FREUD, S. Instincts and their vicissitudes. In: *Collected Papers, IV*. London: Hogarth Press, 1950. (Pp. 60-80.)
13. ———. The Ego and the Id. London: Hogarth Press, 1927.
14. FREUD, A. The Ego and Mechanisms of Defense. New York: Internat. Univ. Press, 1946.
15. FROMM, E. Escape from Freedom. New York: Farrar & Rinehart, 1941.
16. GLUCK, M. R. Rorschach content and hostile behavior. *J. Consult. Psychol.,* 1955, **19**, 475-478.
17. ———. The relationship between hostility in the *TAT* and behavioral hostility. *J. Proj. Tech.,* 1955, **19**, 21-26.
18. HOYT, D. T., & MAGOON, T. M. A validation study of the Taylor manifest anxiety scale. *J. Clin. Psychol.,* 1954, **10**, 357-360.
19. LEWIN, K. Field Theory in Social Science. New York: Harper, 1950.
20. LINDZEY, G., & TEJESY, C. Thematic Apperception Test: Studies of aggression in relation to measurement of overt and covert behavior. *Amer. J. Orthopsychiat.,* 1956, **26**, 567-576.
21. MURRAY, H. A. Exploration in Personality. New York: Oxford Univ. Press, 1938.
22. ROGERS, C. Client Centered Therapy. Boston: Houghton Mifflin, 1951.
23. ROSENZWEIG, S. An outline of frustration theory. In J. McV. Hunt (*Ed.*), *Personality and the Behavior Disorders.* New York: Ronald Pres, 1944. (Pp. 370-388.)
24. SIEGEL, M. S. The relationship of hostility to authoritarianism. *J. Abn. & Soc. Psychol.,* 1956, **52**, 369-372.
25. SROLE, L. Social integration and certain corollaries. *Amer. Sociol. Rev.,* 1956, **21**, 709-716.

26. TAYLOR, J. A personality scale of manifest anxiety. *J. Abn. & Soc. Psychol.*, 1953, **48**, 285-290.

27. ———. The relationship of anxiety to the conditioned eyelid response. *J. Exp. Psychol.*, 1951, **41**, 81-92.

28. WALTERS, R. H., & ZAKS, M. S. Validation studies of an aggression scale. *J. of Psychol.*, 1959, **47**, ———.

Department of Neurology & Psychiatry
Northwestern University Medical School
303 East Chicago Avenue
Chicago 11, Illinois

Department of Psychology
University of Toronto
Toronto 5, Ontario, Canada

Published as a separate and in *The Journal of Psychology*, 1959, **47**, 209-218.

VALIDATION STUDIES OF AN AGGRESSION SCALE*

Department of Psychology, University of Toronto, and Department of Neurology and Psychiatry, Northwestern University Medical School

RICHARD H. WALTERS AND MISHA S. ZAKS[1]

A. INTRODUCTION

The procedure of basing personality scales on clinical categories has one grave disadvantage. The psychologist is forced initially to assess behavior in terms of categories that apply meaningfully only to deviations of an extreme kind. Dissatisfaction with this kind of assessment then leads the psychologist to translate scores on clinical scales into terms that are applicable also to normal behavior. At this point, however, the "objective" character of the scales is diminished, since it is unlikely that all psychologists will agree in their translations, and one of the major advantages of the empirically developed inventory is thus lost.

For example, in Shneidman's *Thematic Test Analysis* (4), John Doe is described, on the basis of his *MMPI* profile, as being anxious, self-depreciating, lacking in self-assurance, and hostile. This part of the description of John Doe's behavior is clearly inferential, since the *MMPI* scales allow for description only in such terms as hysterical, hypochondriacal, psychopathic, schizophrenic and so on. John Doe is a psychiatric patient, yet even in this case the psychologist attempts to describe behavior in terms which are applicable from time to time to almost any individual, normal or otherwise.

Interpretation of scores on scales of the *MMPI* is admittedly difficult. The difficulty increases when the psychologist is dealing with an examinee who is not a psychiatric patient, since it is almost impossible to interpret in a meaningful fashion a profile that consists entirely of moderate and low scores. There has been a proliferation of additional scales for the *MMPI*, some of which, notably the Manifest Anxiety Scale (*MAS*) developed by Taylor (5), have to some extent remedied this defect. However, these scales

---

*Received in the Editorial Office on December 11, 1958, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The authors wish to express appreciation of the assistance given by Harry S. Silverman of the University of Toronto Department of Psychology for his assistance in carrying out the first experiment reported in this paper. The data on which the second experiment was based were obtained by undergraduate students in a psychology course taught by Walters. Detailed mimeographed instructions to *E*s were used to ensure complete uniformity of procedure.

tend to be of an *ad hoc* type constructed in the interests of a particular piece of research and so, with one or two exceptions, hold little promise of providing a systematic and fairly comprehensive set of scales for the assessment and prediction of behavior.

The aggression scale that was employed in the studies reported in this paper represents the first outcome of an attempt to develop a set of scales to measure classes of responses that occur among the general population as well as among clinical groups. The scale consists of 12 items which are included in an inventory of 160 items. Details of the development of the scale and of the manner in which the inventory is administered are given elsewhere (6).

The inventory consists simply of a mimeographed list of items with which *S* indicates agreement or disagreement by circling the appropriate letter, "A" or "D". No intermediate category is presented; *S* is nevertheless urged to answer every item. For group administration items are read to *S* at about 15-sec. intervals. When *S*s are tested individually and are allowed to proceed by themselves, they are urged to work rapidly.

Not all the items in the inventory are new. Some are taken from existing inventories, though occasionally the wording has been modified. For example, 14 *MAS* items were included: these items have been shown by Buss (1) to be valid for non-student groups. A short manifest anxiety scale was thus available for one of the studies reported below.

B. The Validation Studies

1. *Experiment 1*

This experiment was based on the frustration-aggression hypothesis (2). It was predicted that, if the aggression scale was valid, frustrated *S*s should score significantly higher on the scale than non-frustrated *S*s. A supplementary hypothesis, that males should score higher than females, was based on the assumption that in our society males are permitted to express aggression more freely.

a. Method.

(1). *Subjects.* *S*s were students in the third-year general course in social psychology at the University of Toronto. For this course students are divided into two groups on the basis of the first letter of their last name and are taught at exactly the same hours by two different instructors. For the purposes of the experiment these two groups were regarded as being randomly selected. The section taught by the senior author was selected as the experimental group; the other half of the class served as the control.[2]

---

[2] The authors wish to thank Professor J. D. Ketchum for his co-operation in this experiment.

(2). *Procedure.* The experimental group had previously been told that there would be no term examination. On the day selected for the experiment, the instructor handed out a mimeographed list of eight questions. To answer these required a considerable amount of outside reading which, the instructor knew, the students were having difficulty in covering. The group was told that their first assignment of the term had clearly shown that they were not working hard enough, and that it would after all be necessary to give them a test. They were informed that at the next class meeting they would be asked to answer three of the mimeographed questions in an hour's examination, and that in order to do this they would need to have all the material at their fingertips. Moreover, they were told that, since they would not know which three questions would appear on the examination, they would have to prepare all of them thoroughly. Since a weekend intervened between the experimental session and the day set for the examination, the instructor's decision meant that for many of the students the weekend would be spoiled. The students protested that it was impossible for them all to obtain the use of the books they required in the reserve room of the library and that the reading could not be done. The instructor countered these protests by saying that the reading should have been already done, and that he would brook no excuses. Protests continued, as had been anticipated, so the instructor stated that he would use the remainder of the period to run briefly over the topics dealt with in the questions and so "give the students a break".

At this point, as previously arranged, the instructor's assistant walked into the lecture room with the inventories and made a request to administer them for the purposes of research. The instructor thereupon explained to the class that he had agreed that the inventories should be administered, that they were unfortunate that the request had been made that day, but that he was unable to help that. When protests began again, the instructor, to preserve an air of realism, said that he would delete two of the eight questions, but that this meant only that the students would have to answer competently any of the remaining six. After some brief bargaining about which questions should be deleted, the inventories were distributed and administered in the usual way with the instructor reading the items. Tension was high and some students showed a tendency to giggle at certain of the items. This behavior, as well as any restlessness, immediately met with scathing remarks from the instructor.

The deception was successfully carried through until the inventory had been completed and a personal data form attached to the inventory had been filled in. At this point the instructor wrote a list of adjectives on the board and asked each student to write on the back of the inventory those words that best described his feelings about the examination. Conversation with numerous students, including some not actually in the class, disclosed that no one up to this point suspected that an experiment was in progress, and realization was by no means immediate even when this final request was made. At this stage, the students were told that they had been participating in an experiment and that there would be no examination on the following Monday. While a brief explanation was being made, the teaching assistant collected the inventories.

Meanwhile, the other section of the class was administered the inventory in the standard manner, but without any previous frustration. Apart from the preliminary instigation to aggression, and the severe handling by the instructor of any restlessness on the part of the experimental group, the inventories were presented in exactly the same way to both halves of the class.

b. Results.

Since both sections were fairly large, it was decided to pull at random 50 inventories from each group, 25 from males and 25 from females, and to make the comparison on the basis of this number of cases. Inventories which were incomplete were discarded, as were those of *S*s not born in North America. This process of selection in fact exhausted the inventories of male *S*s in the experimental group and all but one of the inventories of male *S*s in the control group. Inventories of 71 female *S*s, 42 from the experimental group and 29 from the control, were not included in the first 100 cases. They provided, however, the data for a second comparison between the groups.

An analysis of the experimental *S*s' statements about how they felt suggested that the procedure had had its desired effect. The adjectives which were written on the board were as follows: irritated, anxious, perturbed, furious, indifferent, glad. Since *S*s were allowed to write more than one word, the total of the frequencies with which the various words were used does not correspond to the total number of *S*s. Five *S*s said they were glad, 66 that they were irritated or furious, 58 that they were perturbed or worried, and three that they were indifferent. On the basis of this analysis, it was decided to compare the responses of the two groups to the 14 *MAS* items, as well as to the aggression items.

The inventories were thus scored for the number of agreements on the 12 aggression items, and also for the number of *MAS* items scored in the keyed direction. Each *S* thus secured two total scores. Those obtained for the first 100 *S*s were entered in the appropriate cells for two 2 × 2 analyses of variance. Means and *SD*s for the four groups are given in Table 1; results of the analyses of variance are given in Tables 2 and 3.

TABLE 1
MEANS AND *SD*s OF MALE AND FEMALE *S*s UNDER CONDITIONS OF FRUSTRATION AND NON-FRUSTRATION ON A 12-ITEM AGGRESSION SCALE AND ON 14 ITEMS OF THE MANIFEST ANXIETY SCALE
(N = 25 in each group.)

	Aggression scale				Anxiety scale			
	Male		Female		Male		Female	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Non-frustrated	3.04	1.71	2.68	1.22	4.88	2.88	5.16	2.24
Frustrated	4.16	1.71	3.52	1.45	5.20	2.51	5.44	2.33

TABLE 2
ANALYSIS OF VARIANCE OF AGGRESSION SCORES
($N = 25$ in each group.)

Source of variance	*df*	Variance estimate	*F*	*p*
Between sexes	1	6.250	2.543	>.05
Between conditions	1	24.010	9.768	<.01
Interaction	1	0.490	0.199	>.05
Within	96	2.458		

TABLE 3
ANALYSIS OF VARIANCE OF ANXIETY SCORES
($N = 25$ in each group.)

Source of variance	*df*	Variance estimate	*F*	*p*
Between sexes	1	1.690	0.259	>.05
Between conditions	1	2.250	0.345	>.05
Interaction	1	0.010	0.002	>.05
Within	96	6.522		

Frustrated *S*s scored significantly higher than non-frustrated *S*s on the aggression scale, thus confirming the main hypothesis. Male *S*s in both the experimental and the control group scored higher than female *S*s, but the difference between sexes was not large enough to be significant. One reason for the failure to confirm the second hypothesis may be that students were used as *S*s; with a non-college group a different result might have been obtained. On the short *MAS* scale there was no difference either between the frustrated and non-frustrated *S*s, or between males and females.

The mean aggression scores of the 42 frustrated and 29 non-frustrated females not included in the first analysis were compared by means of a *t*-test. Frustrated *S*s again scored higher, the difference once more being significant at the .01 level (Table 4).

TABLE 4
AGGRESSION SCALE SCORES OF FEMALE *S*s NOT INCLUDED IN ORIGINAL ANALYSIS

Non-frustrated Mean	Non-frustrated *SD*	Frustrated Mean	Frustrated *SD*	*t*	*p*
5.28	3.10	6.45	2.37	2.38	$<.05$
($N = 29$)		($N = 42$)			

2. *Experiment 2*

Two objections could be made to the first experimental validation of the aggression scale. In the first place, the experiment assumed that the frustration that was induced would result in aggression. Secondly, even if this assumption was justified, the experiment had shown only that the scale was sensitive to situationally-produced aggression. It provided no evidence that it could measure a more or less permanent behavioral characteristic. A further experiment was consequently undertaken in order to test whether the scale could in fact discriminate between normal *S*s who typically showed much or little aggression.

a. Method.

(1). *Development of a criterion.* One hundred and four students in an undergraduate course in psychology were asked to write down six attributes or kinds of behavior that they associated with the term "an aggressive individual." Their responses were tabulated and classified in broad categories in such a way that synonyms or nearly equivalent phrases fell into a single category. Some of the categories involved behavior that is ordinarily approved, e.g., "determined", "energetic"; the remainder involved behavior that is ordinarily disapproved, e.g., "ruthless", "domineering". It was evident that no one description of an aggressive person could be formulated on the basis of the responses. However, when clearly irrelevant and evaluative responses, e.g. "small in stature", "neurotic", were ignored, two descriptions could be drawn up, one embodying the most frequently used "positive" categories, the other the most frequently used "negative" categories. Only the latter description is relevant to the present study.

> There are some people who are ruthless and inconsiderate of others in pursuing their aims. They assert themselves in a domineering manner, and try to impose their ideas on other people. They are loud-mouthed and argumentative, and defy the standards of their own social group.

The description achieved its final form only after being modified as a result of discussion with the class of students that had provided the basic data. In this form it was accepted by the majority of students in a discussion group as a satisfactory description of a negatively aggressive person.

(2). *Subjects.* *S*s all belonged to social groups that allowed fairly close and continuous association among their members. Three adolescent and five young adult groups, consisting in all of 83 male and 33 female members, were involved. The adolescents were members of clubs of the Toronto *YHA*; the young adult groups consisted mainly of University of Toronto students. The nature and composition of the groups are given in Table 5.

(3). *Procedure.* *E* first made out a set of cards, each bearing the name and code number of one of the members of a group. If *E* was a member of the group he was testing, his name was omitted from the cards.

Each *S* was individually tested. *E* spread the cards out in front of *S* and pointed out that his own name was included. *E* said:

> You notice I have a card for you as well as for every other member of this house (club, fraternity). I am going to ask you to make some judgments, and I want you to try to judge yourself as impartially as you judge anybody else in your house (club, fraternity). Don't be afraid to reveal your honest opinion of yourself.

S was now given the description of an aggressive person, and was asked to pick out the .... (approximately 30 per cent) persons in his group who seemed most to fit the description. He was asked to place the cards bearing the names of these persons in a pile at one end of the table or desk. *S* was next asked to pick out the .... (approximately 30 per cent) persons who least fitted the description and to place the cards bearing their names at the other end of the table. *S* was then asked to pick out, from the first of the two piles he had selected, the .... (approximately 10 per cent) of the persons who most fitted the description. These *E* rated "5" on a sheet bearing the code numbers of the persons being rated above the columns and the code numbers of the raters beside the rows. The remaining persons in this first pile were rated "4". *S* was now asked to pick out from the second pile the .... (approximately 10 per cent) persons who least fitted the description. These were rated "1" by *E*: the remaining persons in this pile were rated "2". *E* finally filled in the remaining spaces in the row with ratings of "3". Thus each *S* in effect rated all members of his group, including himself, on aggression in such a way that 10 per cent of the group received a rating of "5", 20 per cent of the group received a rating of "4", 30 per cent of the group received a rating of "3", 20 per cent received a rating of "2", and 10 per cent a rating of "1".

All raters also filled out an inventory. Thus the aggression score of each rater was available. Approximately half the *S*s filled out the inventory before rating their peers for aggression; the other half filled out the inventory after making their rating. Code numbers were used on the inventories as well as on the rating sheet, and *S*s were assured of the confidential nature of their responses. *E*s were careful to explain they could not themselves interpret the inventory and that they did not know what it measured. The total period of testing involved at least two sets of ratings from each *S*, the administration of the inventory and a sociometric test. The results of the additional procedures, however, have no bearing on the present study.

Each *E* set out to obtain ratings by every member of his group on every other member, including *S* himself. A few *S*s were resistant, and were not pressed into coöperation; others were not available at the times when their ratings might have been secured. Ratings were, however, obtained from 106 of a possible 116 raters (Table 5).

b. Results. In order to estimate the reliability of ratings, these were transformed into ranks and Kendall's coefficient of concordance (3) was calculated. The coefficients ranged from .21 to .72 after correction for ties; all were significant at the .001 level (Table 5).

The ratings given to the members of each group were summated, and the totals were ranked. The inventories of the highest ranking and of the lowest

TABLE 5
RELIABILITY OF RATINGS OF AGGRESSION

Nature of group	No. of *S*s rated Male	 Female	No. of raters Male	 Female	*W* (Corrected for ties)
Adolescent Club	6	5	6	5	.30*
Adolescent Club	19	—	13	—	.23
Adolescent Club	—	10	—	9	.49
College Class	—	11	—	11	.72
College Class	11	—	11	—	.55
College Residence	29	—	26	—	.38
College Residence	11	—	11	—	.55
Boarding House	7	7	7	7	.21

*All coefficients are significant at the .001 level.

TABLE 6
AGGRESSION SCALE SCORES OF *S*s RATED HIGH AND LOW ON AGGRESSION
($N = 32$ in each group.)

High rating Mean	 *SD*	Low rating Mean	 *SD*	*t*	*p*
4.25	1.90	2.81	1.86	3.01	$<.01$

ranking third of each group were scored on the aggression scale. Results for the eight groups were then pooled to give aggression scale scores for 32 *S*s rated high in aggression and for 32 *S*s rated low in aggression. Since the same number of high-aggression and low-aggression *S*s was drawn from each subgroup, the aggression groups were balanced in respect to both sex and age.

The means and *SD*s of the aggression scale scores for high-rating and low-rating *S*s are given in Table 6. A *t*-test of the significance of the difference between these means yielded a *t*-ratio of 3.01 ($p < .01$). The result leaves little doubt that the 12 aggression items constitute a valid scale for the measurement of aggression, as defined for this study.

C. SUMMARY AND CONCLUSIONS

Two attempts to validate an aggression scale for normal *S*s were carried out. In the first experiment the inventory was administered to two groups of students. One group was administered the inventory under standard conditions; the other group was frustrated immediately before the inventory was administered to them. Frustrated *S*s scored significantly higher on the aggression scale than did non-frustrated *S*s. Males scored somewhat higher than females, but the difference between the sexes was not a significant one.

In a second experiment eight groups of *S*s were employed. Each *S* rated every member of his group, including himself, for the extent to which he displayed aggression in a socially unacceptable manner. Ratings for each *S* were then totalled. On the basis of these totals *S*s were ranked within their own group for the extent to which they showed aggression. The highest ranking third and the lowest ranking third of each group were regarded, respectively, as high-aggression and low-aggression *S*s. The results from the eight groups were pooled to produce a high-aggression group and a low-aggression group, each of 32 *S*s. Since each rater had already filled in the 160-item inventory, a comparison of the aggression scale scores of the high-aggression and low-aggression *S*s was now possible. *S*s who were rated as high in aggression scored significantly higher on the aggression scale than did *S*s who were rated low in aggression.

There seems little doubt that the 12-item scale is a valid measure of aggression. Not only does it discriminate pathologically aggressive individuals from normal *S*s, but it is also capable of discriminating between *S*s in the normal range. Further refinement may provide a useful instrument for the selection of *S*s for research into the correlates of aggression.

REFERENCES

1. Buss, A. H. A follow-up item analysis of the Taylor Anxiety Scale. *J. Clin. Psychol.,* 1956, **11**, 409-410.
2. Dollard, J., Doob, L. W., Miller, N. E., Mowrer, O. H., & Sears, R. R. Frustration and Aggression. New Haven: Yale Univ. Press, 1939.
3. Kendall, M. G. Rank Correlation Methods. London: Griffin, 1948.
4. Shneidman, E. S. Thematic Test Analysis. New York: Grune & Stratton, 1951.
5. Taylor, J. A. A personality scale of manifest anxiety. *J. Abn. & Soc. Psychol.,* 1953, **48**, 285-290.
6. Zaks, M. S., & Walters, R. H. Preliminary steps in the development of a scale for the measurement of aggression. (To be published.)

Department of Psychology
University of Toronto
Toronto 5, Ontario, Canada

Department of Neurology & Psychiatry
Northwestern University Medical School
303 East Chicago Avenue
Chicago 11, Illinois

Published as a separate and in *The Journal of Psychology*, 1959, **47**, 219-222.

THE EFFECT OF VARYING INTRAMUSCULAR DOSAGES OF ATROPINE AND BANTHINE ON THE GALVANIC SKIN RESPONSE*

Department of Psychology, University of California, Los Angeles

DANIEL J. PERRY, GEORGE E. MOUNT, AND CHESTER D. HULL

Recently studies have been conducted to determine the effects of certain drugs on the measurement of the Galvanic Skin Response (*GSR*). The results of these studies have indicated that this phenomenon can be standardized sufficiently for qualitative and quantitative analysis of drugs acting on the sweat glands (2-5). Also it has been possible to obtain information for the use of these drugs as means of manipulating physiological parameters in the experimental analysis of *GSR* measurements. One of the most recent investigations in this series has indicated that the effects of different anticholinergic drugs administered orally could be compared quantitatively (6). The present study has been designed to see if intramuscularly administered anticholinergic drugs can be evaluated in the same manner.

Ten normal white male subjects ranging in age from 18-31 were studied. Each individual received intramuscular injections of 1/150 and 1/100 grains atropine sulfate, 2 and 4 mg. Banthine (methantheline bromide), and a placebo. The diluent for the drugs and the placebo was 0.5 cc distilled water. The order and sequence of the test substances was varied. Individual tests were made in the morning at least two hours after breakfast and completed before lunch. Two or more days elapsed between successive tests which were performed by the same investigators in the same room.

The *GSR* level of the palms of each subject was recorded for each test for 60 minutes using the previously reported technique (4). Calibrations were made in the beginning, middle, and end of each test run. The high and low resistance levels of each 3¾ minute section of record were determined by linear interpolation between the calibrations for each record. These values were converted into the unit micromhos$^{1/2}$ (5). The data for all subjects were averaged separately for each of the five experimental conditions. The results of these computations involved some variation in initial level for the

---

*Accepted for publication by Roy M. Dorcus of the Editorial Board, received in the Editorial Office on December 12, 1958, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

different drug conditions as shown in Table 1. In order to bring all drug records to the same initial level, corrections were made in the following way.

The "within drugs" regression on the initial values was determined at the 4th, 8th, and final time periods (about 15, 30, and 60 minutes from injection) for each experimental condition. The initial values and regression coefficients used in making the corrections are listed in Table 1. Linear interpolations were used between time periods and corrections computed from the formula $Y'_{at} = Y_{at} - C_t\,(X_{ao} - \overline{X}_{ao})$, (where Y_{at} is the obtained average value for a drug condition at a given time period; $(X_{ao} - \overline{X}_{ao})$ is the deviation from the mean of the uncorrected initial values for the drug condition; C_t is the appropriate correction value, and Y'_{at} is the corrected value). Corrected curves are plotted in Figure 1.

TABLE 1
CURVE CORRECTION DATA

A. Uncorrected initial *GSR* levels expressed in Micromhos$^{1/2}$

Placebo	1/150 Atropine	1/100 Atropine	2 mg. Banthine	4 mg. Banthine
6.1	5.6	5.1	5.9	6.5

B. Regression coefficients used to correct for differences in initial level

4th Time Period	8th Time Period	16th Time Period
0.59	0.39	0.23

The Banthine curves are lower than the other three beginning about eight minutes after injection. This distinction is maintained for the remainder of the experiment. The placebo and atropine curves follow a similar course and tend to separate in the last 15 minutes. At the end of the test period the placebo shows the least effect in lowering the *GSR* curve level followed in order by the two atropine and then the two Banthine curves. The larger dose of both atropine and Banthine ends at a slightly lower *GSR* level (greater resistance) than the corresponding smaller dose.

The statistical significance of the differences for the 4th, 8th, and final time periods were obtained using the error term[1] derived from a treatment by subject analysis of covariance adapted from Lindquist (1). Final computations were made and checked on the Bendix G15 digital electronic analyzer according to a program developed by Mount. The treatment differences are significant at the 5 per cent level in the 4th time period about 15 minutes following the injections (F = 3.67; F 5% df 4/30 = 2.92). They become highly significant in the 8th time period about 30 minutes following injection (F = 4.81; F 1% df 4/30 = 4.51) and remain so in the final time period (F = 13.8; F 1% df 4/30 = 4.51). The differences for 1 per cent significance in the 8th and final time periods are shown in Figure 1.

---

[1] The corrected treatment by subject error term was used. This term is only slightly less than the uncorrected term for a treatment by subject analysis of variance and the primary value of the covariance analysis is in the correction for the obtained treatment effects.

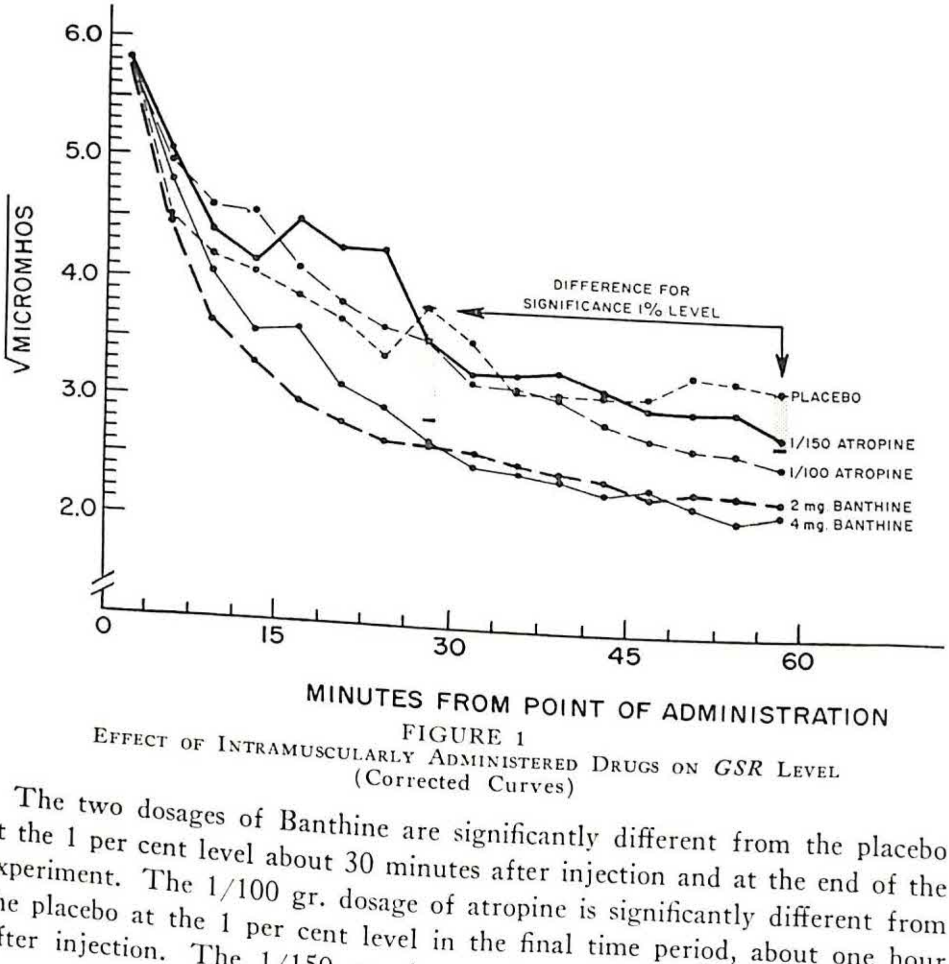


FIGURE 1
Effect of Intramuscularly Administered Drugs on *GSR* Level
(Corrected Curves)

The two dosages of Banthine are significantly different from the placebo at the 1 per cent level about 30 minutes after injection and at the end of the experiment. The 1/100 gr. dosage of atropine is significantly different from the placebo at the 1 per cent level in the final time period, about one hour after injection. The 1/150 gr. dosage of atropine is significant at the 5 per cent level in the final time period. The two dosages of Banthine are significantly different from the 1/150 gr. dosage of atropine at the 1 per cent level at the end of the experiment. Other comparisons can be made from the data of Figure 1.

The results of the present study are similar to those previously reported with orally administered Banthine, Pro-Banthine, and Prantal (5). Regardless of whether the anticholinergic drug is administered orally or intramuscularly, the *GSR* level 45-60 minutes later is significantly lower (decreased sweating) than the placebo. These findings suggest that the *GSR* technique used in these studies may be used to compare dose response effects of different anticholinergic drugs.

REFERENCES

1. LINDQUIST, E. F. Design and Analysis of Experiments in Psychology and Education. Houghton-Mifflin, 1953.
2. PERRY, D. J., & MOUNT, G. E. A comparison of the effect of atropine and placebo on the galvanic skin resistance. *J. Invest. Dermatol.*, 1954, **22**, 497-501.
3. PERRY, D. J., MOUNT, G. E., & HULL, C. D. Effect of drugs on galvanic skin resistance: Comparison of methantheline, propantheline, and a placebo. *A.M.A. Arch. Dermatol.*, 1955, **71**, 476-477.
4. PERRY, D. J., & MOUNT, G. E. Effect of drugs on galvanic skin response level: A study in sympathectomized human subjects. *A.M.A. Arch. Dermatol.*, 1955, **72**, 144-152.
5. PERRY, D. J., MOUNT, G. E., HULL, C. D., & ZEILENGA, R. H. Effect of order of drug administration on the galvanic skin resistance in human subjects. *J. Invest. Dermatol.*, 1955, **25**, 179-185.
6. PERRY, D. J., MOUNT, G. E., & BROWNE, B. W. The effect of varying oral dosages of banthine, pro-banthine, and prantal on the galvanic skin response. *J. Invest. Dermatol.*, 1957, **28**, 239-242.

Department of Psychology
University of California
Los Angeles 24, California

Published as a separate and in *The Journal of Psychology*, 1959, **47**, 223-230.

TWO-YEAR CHANGES ON THE ACE BY STUDENTS IN A COLLEGE OF ART AND ARCHITECTURE*

Rhode Island School of Design

FRANK J. DEIGNAN

A. RATIONALE OF THE STUDY

That liberal arts or teachers college students may be expected to show improvement in intellectual competence throughout their undergraduate years, as measured by such instruments as the college editions of the American Council on Education Psychological Examination or the Ohio State University Psychological Examination, has been demonstrated by the results of several investigations conducted during the past generation. No study of the performance of art students in this area seems to have been done.

The increased interest shown by American educators in the early identification of high school students whose scholastic aptitude and secondary school records indicate that they should consider college-level education is evidenced by the fact that the National Merit Scholarship Corporation administered its test to juniors in high school in April, 1958, and plans to continue this practice, and that the College Entrance Examination Board has of recent years been encouraging high school juniors to take the *CEEB* tests to familiarize themselves with this sort of examination and to reduce emotional tensions when the college board tests are taken in the senior year for the purpose of getting into college. The testing and counseling effects of the recently enacted National Defense Education Act will doubtless multiply this interest. Related to the early identification of intellectually well-endowed students is the accumulation of evidence on the educational potentialities of those high school graduates whose scholastic aptitude for college study is marginal, but who possess other aptitudes that warrant cultivation.

Professional schools of music, architecture, design, nursing, pharmacy, engineering, and possibly teaching, whose curricula are strongly influenced by state certification requirements or by national accrediting standards, place great emphasis in the selection of their students on specialized aptitudes, motivational and personality factors, as well as on high school records and scores on a test of general scholastic aptitude. And although Rhode Island

---

*Received in the Editorial Office on December 12, 1958, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

TABLE 1

SUMMARY OF STUDIES ON CHANGES IN SCHOLASTIC APTITUDE OF COLLEGE STUDENTS AT VARIOUS INTERVALS IN THEIR UNDERGRADUATE EDUCATION

Investigator	Numbers of subjects	Year reported	Test used	Class standing of subjects Initial	Retest	*N*'s	Percentile ranks in Standardization group of total mean raw scores Initial test	Retest	Locale of study
Hartson (4)	403	1936	OSUPE	Fresh	Senior		73	84	Oberlin College
Flory (2)	74	1940	ACE	Fresh	Senior		73	81	Lawrence College
Hunter (5)	276	1942	ACE	Fresh	Fresh*	92	48	71	Converse College
				Fresh	Soph	87	49	73	
				Fresh	Junior	54	62	88	
				Fresh	Senior	43	51	82	
Barnes (1)	105	1943	ACE	Fresh	Soph		59**	84**	College Lib. Arts, Univ. of Illinois
Silvey (6)	517	1951	ACE	Fresh	Soph		44	72	Iowa State Teachers College

*End-of-year freshmen.
**Median raw scores reported.

School of Design may be an exception to the college admissions situation in other American colleges, the number of applicants who completed all of the admissions requirements and whose applications were considered by the Admissions Committee of this college of design has risen from 450 candidates for 150 places in 1953 to 918 applicants for a freshman class of 200 in 1958. Thus, as the number of applications for admission soar it becomes increasingly essential to distinguish between those applicants who have sufficient scholastic aptitude to earn a degree from an accredited college in the northeastern section of the United States from those talented art students who should be encouraged to get further training in a non-degree-granting art school.

Other investigators have reported that statistically significant gains on the *ACE*, or similar tests, may be expected during the college years. Hartson (4), Flory (2), and Silvey (6) all report that those students whose initial scores on the *ACE* placed them at the bottom half or even the lowest quartile of the distribution showed the greatest gains upon re-testing after intervals of two to four years. Table 1 summarizes the data published in a sampling of these studies done over a period of more than 20 years. Although it is possible to equate scores earned on all forms of the college editions of the *ACE* published since 1940 from tables of norms prepared by the publishers, previous investigators have routinely reported the percentile ranks in the standardization group at which the mean raw scores of their subjects would place them. This practice has been followed in Table 1, simplifying a comparison of gains on the *ACE* made by students in various colleges. There is a rather interesting consistency in the level of retest percentile ranks on these tests, regardless of whether it is end-of-year freshmen, sophomores, or seniors who are being retested. Hunter (5) found, in re-administering the *ACE* to the members of each of four undergraduate classes at a small liberal arts college for women in South Carolina, that three-fourths of the total intellectual growth had taken place by the end of the sophomore year. In view of the experiences of other investigators with intellectual growth of college students, and the importance of not summarily eliminating from consideration those otherwise well-qualified applicants for design or architectural education who do poorly on the *ACE*, the administration decided to determine changes in scores on the *ACE* during the first two years by students at Rhode Island School of Design.

B. Design of the Study

The American Council on Education Psychological Examination, 1952 College Edition, was administered during freshman orientation week in September, 1956, to all members of the incoming freshman class and to all transfer students, a total of 230, as one of a battery of tests regularly administered to new students at Rhode Island School of Design. The other tests administered were the Allport, Vernon, Lindzey Study of Values, the Gordon Personal Profile, and the Thurstone Interest Schedule. All tests were administered at one session, the *ACE* being the first in the series. The investigator, dean of students at the college, was in charge of the testing, assisted by two faculty members and six upper division student proctors. The re-testing was done almost exactly 20 months later, during the final month of the students' sophomore year. One hundred eighty of the original 230 were present at the second testing session, the remainder having withdrawn or been dropped from college. On the assumption that there had been negligible carryover in learning from the original test, the 1952 college edition of the *ACE* was used again. Identical testing conditions, as to physical setting, lighting, and test administration, were maintained for both sessions. The *ACE* was administered in May, 1958, as the first in a two-test battery, the other test being the single booklet edition of the Coöperative English Test, Higher Level. Separate answer sheets were machine-scored and hand-checked for both administrations. Means and standard deviations were computed separately for men, women, and for the group as a whole. Product-moment coefficients of correlation were determined between subjects' *Q*, *L*, and Total scores on their performances on the *ACE*. Using Garrett's (3, p. 218) formula for computing standard errors of the differences between correlated means, significances of the differences were determined. To test the gain hypothesis of those subjects whose original scores placed them in the lowest quartile of the distribution, Q_1 and Q_3 of the initial test were determined. Test-retest correlations were computed between the scores of those who fell below Q_1, those within the interquartile range, and those who scored above Q_3 on the original testing.

C. Results and Discussion

Table 2 presents in tabular form the most important findings of this study of changes on the *ACE* scores of art and architecture students after two years of college education. With the exceptions noted by the absence of superscripts in Table 2, all changes in scores from the initial test to the retest were gains significant beyond the .01 level of confidence. The female subjects earned higher *L*-scores and higher Total scores than did the males on both administrations of the test. On both the initial test and the retest the Linguistic mean raw scores of the women were significantly better than those of the men at the .01 level of confidence. On the mean raw scores for the

TABLE 2

A COMPARISON OF SCORES ON THE 1952 COLLEGE EDITION OF THE *ACE* MADE BY STUDENTS OF ART AND ARCHITECTURE AS BEGINNING FRESHMEN AND AS END-OF-YEAR SOPHOMORES

Part of *ACE*	Group	Number of subjects	Raw score means Fresh	Raw score means Soph	Gains	Percentile ranks in norm group of means Fresh	Percentile ranks in norm group of means Soph	Standard deviations Fresh	Standard deviations Soph	Test-Retest *r*'s
Total scores	Total	180	116.16	126.25	10.09*	65	78	18.05	17.90	.80
	Men	101	113.15	123.95	10.80*	61	76	17.39	16.91	.75
	Women	79	119.63	128.62	8.99*	70	81	18.34	18.95	.84
Q-Scores	Total	180	44.36	46.20	1.84	60	66	9.37	9.20	.70
	Men	101	44.50	47.55	3.05*	60	73	9.50	8.27	.67
	Women	79	43.73	44.31	.58	60	60	9.00	9.91	.75
L-Scores	Total	180	71.80	80.05	8.25*	66	80	12.95	13.36	.84
	Men	101	68.65	76.40	7.75*	59	73	12.05	11.53	.83
	Women	79	75.90	84.31	8.41*	73	86	12.95	13.35	.84

*Gains significant beyond the .01 level of confidence.

Totals, however, the degree of confidence that the women did better than the men is somewhat lower; there are 95 chances in 100 that the greater Total mean raw scores of the women on both the first and second tests represents superior performances by the women and are not attributable to chance differences. The mean raw *Q*-scores of the men exceed those of the women on both test administrations. Although the difference between mean *Q*-scores of men and women on the initial test is negligible, that on the retest is significantly greater for the men at the .05 level of confidence.

These changes in scores on the *ACE* after two years in college differ somewhat from those reported by Silvey (6), who found that his male subjects had higher mean raw scores on all parts of the *ACE* than did his women subjects on both administrations of the test. The School of Design data, however, conform quite well with those reported by Flory (2, p. 447) ". . . the women exceeded the men in percentile rank at both the freshman and senior levels. . ." The superiority of School of Design women on *ACE* total scores may be due to the fact that with numbers of applications from women greatly exceeding those from men the admissions committee can be much more selective in its acceptance of women candidates. A large proportion, about 40 per cent, of the male sample at Rhode Island School of Design is majoring in architecture, landscape architecture, and machine design, all of which require several courses in higher mathematics. Thus, the greater gains in *Q*-score means made by the male subjects in this study is consonant with Hartson's (4, p. 490) conclusion, ". . . improvement in intellectual abilities continues during the college period in the fields in which that experience provides exercise. Continued maturation is not a general process of unfoldment of inner capacities, which occurs independently of particular training. . ."

The results of an analysis of the retest performance of those subjects who on the first test were in the top quartile, the middle of the distribution, or the lowest quartile, are presented in Table 3. All groups made significant gains in total mean raw scores on the retest. The greater magnitude of the mean raw score gains, and in percentile rank position of the retest mean of the lowest quartile subjects agrees with the findings of other investigators that those on the bottom of the original distribution tend to make greater gross gains, and greater gains in the percentile rank position of the mean retest scores than do those subjects who do better on the original test. Despite the evidence of improvement in scholastic aptitude, however, there remains a disturbing discrepancy in intellectual ability between those students at the lowest quartile and those in the top quartile of the group. Is it possible for

TABLE 3

RELATIVE GAINS ON *ACE*-TOTAL SCORES AFTER TWO YEARS IN COLLEGE OF STUDENTS INITIALLY IN THE TOP QUARTILE, THE INTERQUARTILE RANGE AND THE LOWEST QUARTILE OF THE DISTRIBUTION

Group	Number of subjects	Total raw score means		Raw score gains	Standard deviations		Percentile ranks in norm group of means		Test-Retest *r*'s	*t*-values
		Fresh	Soph		Fresh	Soph	Fresh	Soph		
Above Q_3 (127.3)	45	139.11	145.20	6.09	9.47	10.48	90	94	.86	5.49*
Within Interquartile range (103.7 to 127.3)	90	115.94	125.41	9.47	6.44	11.30	65	77	.30	7.89*
Below Q_1 (103.7)	45	95.43	106.71	11.28	8.00	11.74	35	52	.39	8.67*

*Significant beyond the .001 level of confidence.

a faculty to do effective teaching with students who display such wide disparities in intellectual abilities? This is a particularly crucial question in a small institution offering degrees in 12 different areas of art and architecture with the consequence that homogeneous grouping on the basis of intellectual ability is impossible.

D. SUMMARY

The 1952 College Edition of the American Council on Education Psychological Examination was administered to a group of 230 beginning freshmen and transfer students of art and architecture, and re-administered to 180 of the same group who were still in attendance at the conclusion of their sophomore year. Statistically significant gains (.01 level) were made in seven of the nine raw score means which were compared. The gains of the female subjects and of the total population on the *Q*-scores were not significant. Comparison of test-retest performance of subjects whose original scores placed them above Q_3, within the interquartile range, or below Q_1 disclosed that those below Q_1 made greater raw score gains, and achieved greater gains in percentile rank position of the mean raw scores in the standardization group than did those subjects in the other two groups. The difficulties involved in maintaining satisfactory levels of instruction for groups so disparate in scholastic aptitude were emphasized.

REFERENCES

1. BARNES, M. W. Gains on the *ACE* during freshman-sophomore years. *Sch. & Soc.*, 1943, **57**, 250-252.
2. FLORY, C. D. The intellectual growth of college students. *J. Educ. Res.*, 1940, **33**, 443-451.
3. GARRETT, H. E. Statistics in Psychology and Education. New York: Longmans, Green, 1946.
4. HARTSON, L. D. Does college training influence test intelligence. *J. Educ. Psychol.*, 1936, **27**, 481-491.
5. HUNTER, E. C. Changes in scores of college students on the American Council Psychological Examination at yearly intervals during the college course. *J. Educ. Res.*, 1942, **36**, 284-291.
6. SILVEY, H. M. Changes in test scores after two years in college. *Educ. & Psychol. Meas.*, 1951, **11**, 494-502.

Rhode Island School of Design
Providence 3, Rhode Island

Published as a separate and in *The Journal of Psychology*, 1959, **47**, 231-234.

DIFFERENTIATION AND DIFFERENTIAL COUNSEL OF SOME NEUROTIC PERSONALITIES*

Akron Psychological Consulting Center

EDWARD S. KARON AND BERTRAM P. KARON

There has been increasing skepticism about the value of diagnostic categories since they seem to be unrelated to differential treatment. It is not surprising, therefore, that such a prominent psychotherapist as Sullivan (5, pp. 74-77; 6, p. 4) expressed doubts about utility of classical diagnoses, and Rogers (3) even advocates its abandonment for neurotic patients on the basis that the diagnosis does not aid his task. It is thus worthwhile to describe one set of diagnostic categories which we have found to have important implications not only as to the nature of the patient's problem but as to the way it should be handled.

The majority of neurotic patients who have come to our attention we have found separable into two reaction patterns which we will refer to as the pre-schizophrenic and pre-manic-depressive personalities. This is roughly similar to the classical schizothymic vs. cyclothymic, introvert vs. extrovert, intoceptive vs. extraceptive, schizoid vs. cycloid personalities, a distinction featured prominently in the works of Jung (2) and Rorschach (4) among many others. Eysenck (1, pp. 19, 123-152) presents a good survey of some of the literature on this distinction.

The categories described below are based on over seven years experience in an out-patient clinic with more than 150 patients, approximately equally divided between men and women. Age range was from 19-36. For this population it had proven relatively simple to assign the majority of patients to these two categories, and the categories have served to provide an accurate summary of many important features of the patients' personalities. We do not claim that only bi-polar personality patterns exist or that our clinical implications would apply with the psychotic. The terms pre-schizophrenic and pre-manic-depressive are simply labels of personality types that could theoretically operate on any level of adjustment from normal through psychosis.

The pre-schizophrenic, although markedly idealistic, plays the overt rôle

---

*Received in the Editorial Office on December 13, 1958, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

of the cynic. His defense against his problems is withdrawal from people; under stress this reaction may terminate in the severe withdrawal from life that marks psychosis. Characteristically the pre-schizophrenic neurotic shows compulsive behavior, inability or marked difficulty in relating emotionally, and a fear of emotional spontaneity. Although often highly intelligent with a good deal of social "know how" and insight into their problems, they are unable to utilize these abilities except to increase depression. There seems to be continuous conflict between conscious and unconscious goals; either may be dominated by drive toward dominance or submission, activity or passivity.

Typically the personality of the pre-schizophrenic neurotic is segmentalized with limited emotional interaction between areas of functioning. Symptoms tend accordingly to be limited to specific areas of functioning and often become truly psychosomatic with organic involvement.

The pre-manic-depressive category as contrasted with the pre-schizophrenic presents a facade of idealism which does not extend below the surface; inwardly he seems a true cynic with devaluation of even his own codes. Noticeably overemotional, he tends to think in extremes of affect, preferring to love or hate rather than to like or dislike. Although typically deficient in insight into himself and others, he can understand those with similar temperament. He is commonly extremely suggestible but usually develops mechanisms to conceal both these and his overemotional reactions.

The pre-manic-depressive individual reacts wholistically: one maladjusted area taints his entire functioning. Thus these people are much more prone to the blatant symptoms of a generalized hysterical reaction and to changeable symptoms rather than to the development of the fixed and more carefully delineated symptoms of the pre-schizophrenic.

From these characteristics it is easier to predict certain aspects of any psychotherapeutic relationship. The pre-schizophrenic personality is very slow to develop trust in his counselor, or in psychoanalytic terms a workable transference. Once made, the bond can usually withstand the stresses of even intensive depth counseling.

The outwardly idealistic pre-manic-depressive personality seems to form quick bonds, accepts readily any value judgments he can entice from his therapist. The patient's suggestibility and apparent wholistic improvement in the earlier sessions lead the naïve counselor to expect a relatively short cure or drastic amelioration of symptoms. Clinical reality affords disillusionment to the novice: workable transferences are rare and almost unidentifiable.

The pre-schizophrenic personality usually improves or gets worse at a gradual pace. The tendency towards concretely defined psychosomatic ailments seems to act to retard emotional disillusionment. The individual has learned to build walls against even the horrors of psychosis; attempts to smash these defenses usually result in the construction of newer and stouter ones. The most productive approach seems to have been to win slowly the patient's confidence and trust while refusing to allow him to intellectualize. Under such handling he will gradually integrate his conscious and unconscious goals. His pre-existing insight into others and himself can be used as a starting point, and he can be led to extend this recognition of his problems without direct coercion. The therapist has a potent weapon in his awareness of the pre-schizophrenic patient's guarded but very real idealism; recognition of this is a major aid in invoking a reversal of the patient's withdrawal reaction. Since the pre-schizophrenic nature rarely responds wholistically, concentration on a particular symptom is wasteful; unless the patient is stimulated by the therapist to seek re-integration, he will simply revert to a state of perpetually isolated symptoms.

The pre-manic-depressive personality is characteristically rapid in his emotional responses, and this is generally mirrored by rapid changes during therapy. These are generally the cases that improve drastically with short-term directive counseling, short-term non-directive counseling, abbreviated psychoanalysis, therapy from drugs, and diets of bread and water. The one certainty is that they will change. Apparently only when adjustment is extremely good or extremely poor can these people hold to a consistent pattern of functioning. The consistent, sympathetic, understanding therapist that is needed in order to deal with the pre-schizophrenic has to be here replaced by a therapist who himself can follow the patient's varying moods. The false shell of idealism can be effectively dealt with by directly calling attention to the patient's own real guarded cynicism. The patient is not offended as would be the pre-schizophrenic, but rather reacts as if it is a relief that the facade need no longer be maintained. Consistency of treatment beyond this is impractical. Neither a purely directive, nor a purely intellectual, nor a purely non-directive approach has validity. The most consistent improvement is attained by the therapist who is both sensitive to the patient's changing emotional reactions and flexible enough to match his therapeutic technique to the patient's changing needs. Any symptom pattern can be taken as the focus of the therapeutic effort since the patient will extend any gains throughout his whole personality structure.

Typically the pre-schizophrenic is characterized by the necessity for relatively long periods of therapy to effect a cure, although, if the therapist passes over specific symptoms in favor of dealing with the more basic structures, the duration can be effectively shortened. When therapy is successful or even partially successful with such patients there seems to be minimal chance of a regression.

Despite the false hopes engendered by the typical behavior in counseling of the pre-manic-depressive personality, these patients actually do offer the counselor the greatest chance for dramatic improvement and massive symptom dispersal. To improve these cases affords no real difficulty; to cure is another matter. Tapering the patient's treatment with increasing intervals between appointments seems effective here. One might say that they are cured easily and quickly, but one must spend much time and many sessions seeing if the cure really is a cure.

REFERENCES

1. EYSENCK, H. J. Cyclothymia and schizothymia as a dimension of personality: Historical review. *J. Personal.*, 1950, **19**, 123-152.
2. JUNG, C. G. Psychological Types. New York: Harcourt, Brace, 1933.
3. ROGERS, C. R. Client-Centered Therapy. Boston: Houghton, Mifflin, 1951.
4. RORSCHACH, H. Psychodiagnostik. Bern: Huber, 1942.
5. SULLIVAN, H. S. Conceptions of Modern Psychiatry. New York: Norton, 1953.
6. ———. The Interpersonal Theory of Psychiatry. New York: Norton, 1953.

Akron Psychological Consulting Center
788 North Howard Street
Akron 10, Ohio

Published as a separate and in *The Journal of Psychology*, 1959, **47**, 235-245.

THE EFFECTS OF FLOODING AND PHYSICAL SUPPRESSION AS EXTINCTION TECHNIQUES ON AN ANXIETY MOTIVATED AVOIDANCE LOCOMOTOR RESPONSE*

Teachers College, Columbia University

A. Terrence Polin[1]

A. Problem

According to Pavlovian theory, all that is required for extinction to take place is to present the conditioned stimulus without the unconditioned stimulus. However, several studies of avoidance behavior (Estes, 1944; Estes & Skinner, 1941; Masserman, 1943; Sidman, 1955) have reported that a conditioned instrumental response is extinguished very slowly despite the omission of the unconditioned stimulus. The difference in results concerns clinicians in their search for successful methods for extinguishing "unhealthy" avoidance behavior.

Among extinction techniques, blocking by suppressing the occurrence of the conditioned instrumental response while the conditioned stimulus is being presented (Miller, 1948; Mowrer & Lamoreaux, 1942; Page & Hall, 1953; Schoenfeld, 1950; Solomon & Brush, 1956) has been used with little agreement as to its effectiveness. Estes (1944) argues that Pavlovian extinction by physically suppressing the instrumental response may be punitive. In addition to physical suppression, Mowrer & Lamoreaux (1942) suggest that massing of extinction trials (allowing the subject to make the conditioned instrumental responses freely) accelerates avoidance extinction. They (Mowrer & Lamoreaux, 1942, p. 30) generalize that "in any situation in which anxiety is the chief motive, anxiety reduction (reward) is a function . . . of the rate at which the danger signal is presented. From this it follows that massing of trials would be even more effective in producing extinction . . . in the case of avoidance conditioning than in other forms of conditioning."

Most avoidance studies (Solomon & Brush, 1956) on the effectiveness of massing and blocking extinction have used only one cycle of conditioning and

---

*Received in the Editorial Office on January 9, 1959, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This experiment is based on a doctoral dissertation conducted under the direction of Professors I. Lorge, P. M. Symonds, and R. L. Thorndike, and submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Teachers College, Columbia University, in 1958.

extinction. With two cycles of conditioning and extinction the present study investigated the stability and rapidity of extinguishing a shock-avoidance locomotor response by mass trials and physical suppression.

As compared with rest, Pavlovian extinction by physically suppressing the animal's emission of the conditioned instrumental response should result in a diminished overt response strength. But physical suppression of the conditioned instrumental responses is not expected to weaken permanently or eliminate the latent response strengths but merely to repress the avoidance behavior. Since a barrier prevents the animals from making a free avoidance response to the anxiety-laden situation, any "emotional" behavior similar to the initial conditioning experience should arouse the animal to emit the repressed instrumental responses in addition to the newly acquired responses. This should result in more unreinforced avoidance responses as compared with rest. However, as compared with rest and barrier extinction, conditioned avoidance behavior after flooding should be extinguished quickly and permanently since the animal has freely emitted a high number of unreinforced instrumental responses.

B. METHOD

1. *Subjects*

Thirty experimentally naïve, male, albino rats of the Wistar strain were randomly separated into three apparently equivalent groups of 10 animals. For the purpose of identification the groups are referred to as: (*a*) Control, (*b*) Barrier, (*c*) Flooding.

2. *Apparatus*

A modifier Mowrer-Miller box was divided into two equal compartments by a 1½" hurdle made of lucite. Directly above and parallel with the hurdle were two brass rods which were electrically charged by a separate electrical circuit with no appreciable resistance, delivering 75 volts. The hurdle was charged independently to prevent animals from perching on it while in the experimental box. The floor of each compartment consisted of a grid, wired so that it could be charged independently of the other. The shock was 175 volts delivered through a 375,000-ohm resistance in series with the grid. The shock was sufficiently potent to prevent the animals from freezing but not strong enough to cause any permanent physical damage to them. The *CS* was a buzzer suspended beneath the conditioning box.

3. *Design and Procedure*

The experimental procedure of 40 days duration was divided into six periods:

a. Day 1, measuring initial hurdle-crossing.
b. Days 2-11, conditioning (I).
c. Days 12-15, experimental variations Control, Barrier, Flooding.
d. Days 16-28, extinction (I).
e. Days 29-30, re-conditioning (II).
f. Days 31-40, re-extinction (II).

a. To get estimates whether the three groups may be considered equivalent in initial response of hurdle-crossing to the sound of the buzzer, each animal received 20 trials of five seconds of buzzer. The number of hurdle-crossings to the buzzer was recorded. The groups were considered equivalent.

b. Conditioning (I). Following a two-minute acclimatization period, each animal was given 20 trials of avoidance training. Each trial consisted of (*a*) the presentation of buzzer which continued to sound for five seconds if the animal did not cross over the hurdle, (*b*) the onset of the electric shock together with the continuation of the buzzer sound until the animal crossed the hurdle into the "cold" compartment. If the animal responded to the buzzer before the onset of shock by crossing the hurdle into the other compartment, the buzzer was immediately terminated manually. Trials were separated by a constant interval of 30 seconds.

c. Experimental Extinction. In the four days following the tenth and last day of avoidance conditioning (I), the different experimental extinction treatments were used with the three groups. The Control Group received four days of rest in the living cages. The Barrier Group received four days of barrier extinction in which daily the process was as follows: (*a*) two-minute acclimatization period, (*b*) buzzer for fixed period of five seconds regardless of animal's behavior, (*c*) 30 seconds inter-trial interval separating termination of buzzer and onset of buzzer again for the next trial. No shock was administered at any time during the barrier trials. After the two-minute acclimatization period, a sheet of clear lucite was used as a barrier to separate the two compartments. The lucite barrier extended over the hurdle and reached to the ceiling of the box.

The Flooding Group received four days of the following extinction process: (*a*) two-minute acclimatization period, (*b*) onset of buzzer and continuation of buzzer for time period equivalent to 20 buzzer periods (100 seconds) with no inter-trial rest periods. No shock was presented at any time. Thus, the animals were free to cross the low hurdle at any time.

d. Extinction (I). The regular extinction trials after the four days of experimental extinction were similar to the avoidance conditioning trials except that the unconditioned stimulus (shock) was never given, and the buzzer was on for a fixed period of five seconds instead of being terminated when the animal crossed over the hurdle. All animals received 20 trials per day for a period of 13 days.

e. Re-conditioning (II). Two days of avoidance training, identical with the initial avoidance training period, were given after the thirteenth day of regular extinction.

f. Re-extinction (II). After the two-day reconditioning of avoidance, 10 days of regular extinction followed, identical with the previous 13-day period of Extinction (I).

The interpretation of the results is based on the hurdle-crossing responses during extinction and re-extinction.

C. RESULTS

1. *Conditioning Phase I Analysis of Data*

The initial avoidance conditioning data for the three groups shown in Table 1 demonstrate that the three groups conditioned at a similar rate. Analysis of variance of repeated measurements for the three independent groups showed no significant difference among the three conditioning curves. All three groups attained and maintained a high level of avoidance conditioning by the eighth day, averaging 19 avoidance responses out of 20 trials.

TABLE 1
AVOIDANCE JUMPS OF THREE GROUPS OF 10 ANIMALS DURING 10 DAYS OF CONDITIONING: MEAN NUMBER OF JUMPS PER 20 TRIALS PER DAY

Animal groups	Days									
	1	2	3	4	5	6	7	8	9	10
Control	1.7	4.6	8.5	14.5	16.2	17.6	18.6	19.0	19.5	19.9
Barrier	1.5	4.1	8.2	14.4	16.5	17.8	18.5	19.1	19.8	19.7
Flooding	1.7	4.6	8.7	14.9	16.2	17.4	18.7	19.1	19.6	19.8

2. *Measuring Effects of Extinction*

No quantifiable record was made of the "apparent" or incipient avoidance responses by the Barrier animals. Nevertheless, the animals made many attempts to cross over into the adjoining compartment, and much "emotional" behavior was observed during the four days, e.g., excessive defecation, urination, vertical jumping, scratching the walls, and continuous restless exploration.

In the Flooding Group, each animal received daily the equivalent of 20 buzzer trials uninterrupted by an inter-trial interval. During the four days of flooding, the box was unmodified and identical to that used throughout the conditioning phase. The eliciting strength of the 100-second continuous buzzer is the number of hurdle crossings during the four-day period. The Flooding Group made a total of 563 crossings to the buzzer.

3. *Extinction Phase I Analysis*

The effect of the three different experimental treatments was measured by the number of avoidance responses made during the 13-day regular extinction phase I, as shown in Table 2. The immediate consequence of each extinction treatment is indicated by the group differences in resistance to extinction during the first few days of extinction phase I. Barrier operant training which was Pavlovian extinction, affected initially the number of instrumental jumping responses as compared with the Control group. However, the rate of extinction differed substantially; indeed, the two extinction curves crossed about half way through the regular extinction phase. As compared with the Control and Barrier Groups, the Flooding Group extinguished more rapidly. The three groups differ significantly ($p < .01$) in the rate of extinction as shown in Figure 1. However, all three groups were responding at a very low level during the last three days of extinction.

TABLE 2
AVOIDANCE JUMPS OF THREE GROUPS OF 10 ANIMALS DURING 13 DAYS OF EXTINCTION: MEAN NUMBER OF JUMPS PER 20 TRIALS PER DAY

Animal groups	Days												
	1	2	3	4	5	6	7	8	9	10	11	12	13
Control	18.1	15.6	14.0	11.5	9.7	7.1	5.9	5.5	4.9	4.0	3.7	3.0	2.6
Barrier	13.5	11.9	10.6	9.4	8.2	7.1	6.0	5.7	5.6	5.5	5.3	5.0	4.9
Flooding	6.4	5.5	5.0	4.5	4.1	3.8	3.5	3.3	3.0	2.8	2.5	2.2	1.7

4. *Reconditioning Phase II Analysis of Data*

There were no significant differences among the three groups in the rate of re-conditioning (Table 3) in the next two days, and the three groups of animals quickly attained a common high level of avoidance behavior.

TABLE 3
AVOIDANCE JUMPS OF THREE GROUPS OF 10 ANIMALS DURING TWO DAYS OF RECONDITIONING: MEAN NUMBER OF JUMPS PER 20 TRIALS PER DAY

Animal groups	Days	
	1	2
Control	17.9	19.6
Barrier	17.5	19.5
Flooding	17.8	19.6

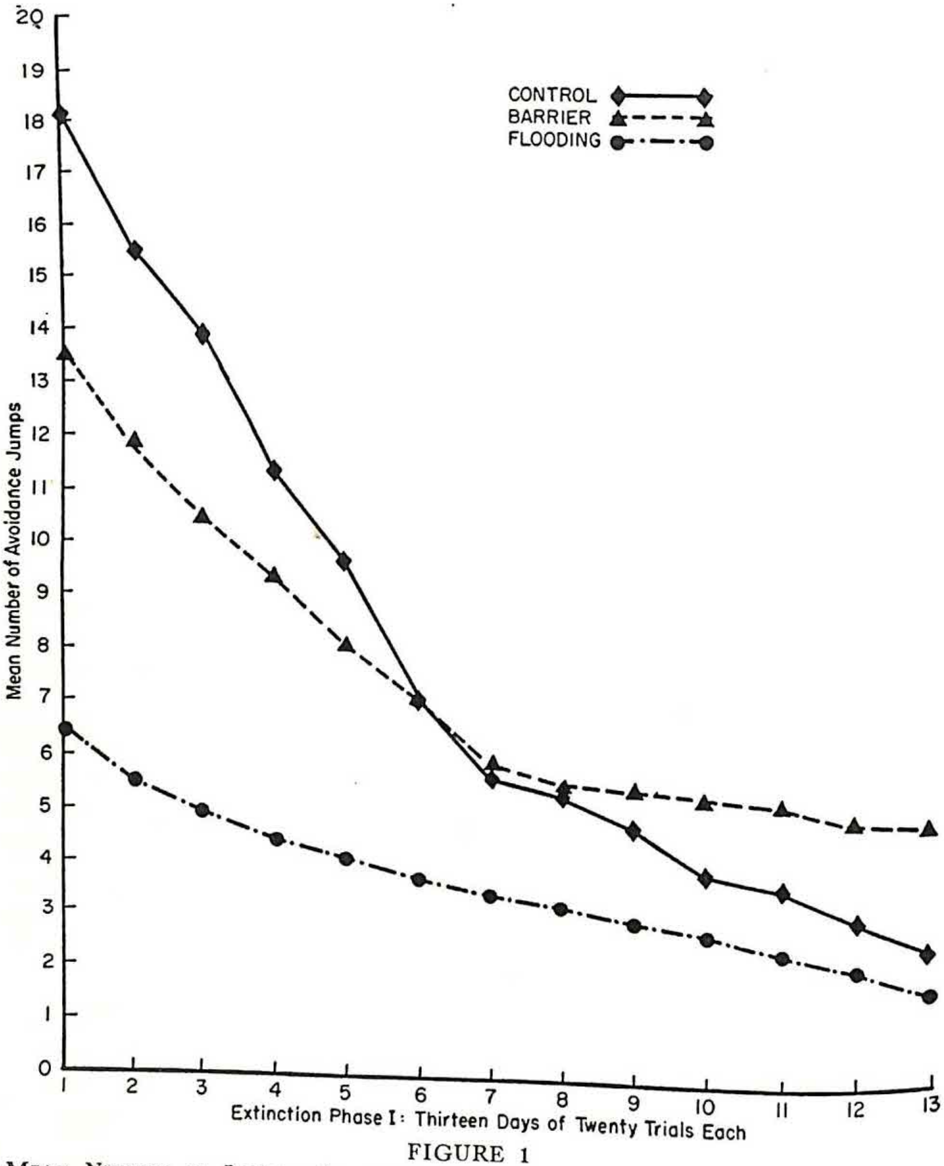


FIGURE 1

MEAN NUMBER OF INITIAL AVOIDANCE JUMPS TO BUZZER BY THREE GROUPS OF 10 SUBJECTS DURING 13 DAYS OF REGULAR EXTINCTION

5. *Re-extinction Phase II Analysis of Data*

The re-extinction data are reported in Table 4. The three curves differ significantly in rate. Although the three groups did recondition at similar rates, they did not re-extinguish at similar rates. The three extinction curves are significantly different ($p < .01$) in height and shape from each other as shown in Figure 2.

TABLE 4
AVOIDANCE JUMPS OF THREE GROUPS OF 10 ANIMALS DURING 10 DAYS OF RE-EXTINCTION: MEAN NUMBER OF JUMPS PER 20 TRIALS PER DAY

Animal groups	Days 1	2	3	4	5	6	7	8	9	10
Control	17.6	16.2	14.2	10.6	9.5	8.1	5.8	4.3	3.7	2.9
Barrier	19.5	18.3	17.7	16.9	16.7	16.1	15.7	15.4	14.7	13.6
Flooding	11.4	7.8	6.2	4.3	3.7	3.2	2.8	2.5	2.1	1.8

In the re-extinction phase, the Barrier Group showed marked resistance to extinction. This would appear to support the generalization that extinction by physically suppressing does not permanently weaken avoidance instrumental behavior. The reconditioning experience resulted in bringing forth a large number of responses during the subsequent extinction.

Since each group's rate of avoidance on the last days of conditioning and of reconditioning was similar, the effects of the three treatments in each extinction phase were measured. The effect of the four days of experimental extinction techniques as compared with the first four days of regular extinction was measured by analyzing the number of *CR*'s made during the first six days of extinction phase I and was compared with the fifth through tenth days of re-extinction. Since significant differences ($p < .01$) were obtained between the two time periods, the effects of the four-day treatment period were statistically different from the first four days of regular re-extinction trials.

For each of the three groups, the hypothesis was rejected that the heights were similar between the two extinction periods. Since the extinction curve of a control group should approximate a typical *S*-shape, it is expected that the comparison of the two extinction periods would differ as to shape. The first six days of extinction phase I plots the top half of the *S* curve. The last six days of the re-extinction phase II curve is the bottom half and is concave up rather than down. Four days of rest or disuse is not equivalent to the same number of days devoted to regular extinction.

Although the two extinction curves of the Flooding Group differed as to height or level of responding, there was no significant difference between the

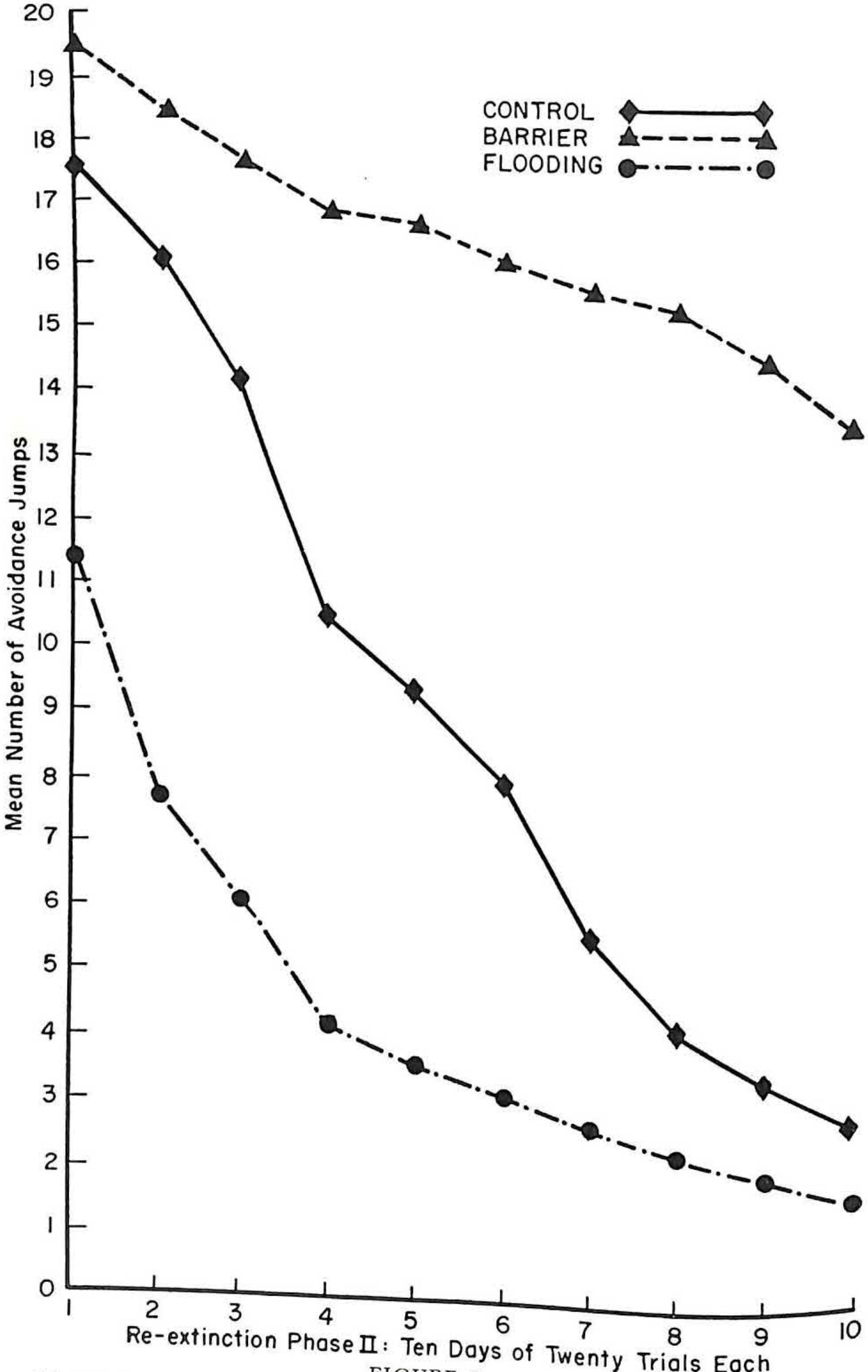


FIGURE 2

Mean Number of Avoidance Jumps of Three Groups of 10 Subjects in a Series of 10 Days of Regular Extinction shape of the curves. The *F* ratio comparing the shape of the Barrier Group's extinction curves is significant at the 5 per cent level but not from the comparable number of regular extinction trials. Reconditioning appears to heighten the resistance of the Barrier Group to regular extinction.

D. Discussion

From the analysis of the data it is reasonable to infer that the three groups of rats were from a common population prior to the onset of the four-day experimental extinction treatments. Experimental extinction of a locomotor response acquired to avoid shock was affected to a significant extent, by the special extinction treatments of flooding and barrier as compared with rest or removal from the experimental environment.

Using the Control Group's extinction phase I data as a benchmark, Pavlovian extinction by blocking the instrumental responses resulted in fewer emitted hurdle-crossings by the Barrier Group during the early days of regular extinction. However, the intersection of the two 13-day curves indicates that the over-all resistance rate of the Barrier Group persisted at a higher rate after six days. There was a significant difference in the rate or level of effectiveness between the two treatments for 13-day extinction. Physical suppression, suggesting punitive properties, has a more flattening effect on the rate of extinction than the rest group. Pavlovian extinction without instrumental extinction does not lead to a rapid extinction rate beyond the initial days. The barrier treatment has a deleterious effect on the extinction of an unnecessary instrumental avoidance response as measured over a specific period.

The difference in number and rate of avoidance responses between the two treatment groups in re-extinction phase II indicates that the instrumental avoidance behavior of the Barrier animals has considerable latent strength. An "emotional" occurrence similar to the previous experimental condition arouses in the organism a fearful state or drive which results in emission of a significant number of repressed instrumental responses. The "therapeutic" value of physical suppression as an extinction technique is questionable; indeed, it may be punitive.

In both extinction phases the Flooding Group extinguished more rapidly than either the Control or Barrier Group. Although the Flooding Group showed rapid extinction during the regular extinction phases, the animals reconditioned as efficiently as the other two groups. The massing of instrumental avoidance responses had no significant effect on reconditioning. It appears that the fear-producing properties of the conditioned stimulus (buzzer, proprioceptive, and tactile) are reduced significantly when the organism is free to emit the avoidance or escape behavior until it has eliminated the fear of the situation.

E. SUMMARY

This study tested the relative effectiveness of extinguishing an instrumental avoidance response by massed and blocked extinction trials. Within the framework of the two-process learning theory, it was argued that applying only Pavlovian extinction by physical blocking of responses represses rather than permanently weakens instrumental avoidance strength.

Three groups of rats, consisting of a control and two experimental groups, were conditioned to avoid electric shock by jumping from one compartment to the adjoining one in a Miller-Mowrer box. After 10 days of avoidance conditioning, during which the animals were given 20 trials per day for 10 days, the three groups reached the required criterion of conditioning. Each group was given four days of experimental extinction training. During these four days, the three groups were given different conditions for extinction. The Flooding Group received daily 100 seconds of continuous stimulus (buzzer) and was free to emit the instrumental avoidance response. The Barrier Group was given 20 Pavlovian extinction trials daily, with a barrier positioned to prevent the animals from moving to the adjoining compartment. The Control Group was isolated in the living cages for the four days.

After the four days of differential extinction procedures, the three groups were given identical treatment for 13 days of extinction. Although the Barrier or Pavlovian extinction Group showed less resistance to extinction in the early stages as compared with the Control Group, after six days of extinction trials the Control Group began making fewer instrumental responses. The Flooding Group exhibited the best extinction results as compared with the other two groups. Immediately after the 13-day test, all animals were given two days of reconditioning. All three groups were conditioned to a high level of avoidance and were not significantly different from each other. Immediately after the two days of reconditioning, the three groups received 10 days of regular re-extinction. The Pavlovian extinction or Barrier Group demonstrated a persistent resistance to extinction. The Flooding Group quickly adapted to the extinction period and showed significantly fewer instrumental responses. The statistical analysis gave support to the following conclusions:

1. Blocking of the instrumental avoidance responses for four days after conditioning resulted in fewer number of emitted hurdle-crossings in contrast to the Control Group. The crossing of the phase I curves for Control and Barrier indicated that the resistance of the Barrier group was somewhat stronger than was hypothesized.

2. Blocking of the instrumental avoidance responses represses but does

not eliminate the latent instrumental strength to respond under similar "anxiety" conditions. Reconditioning of the Barrier Group led to the maintenance of greater resistance to regular extinction as compared with the Control and Flooding Groups.

3. Flooding or free emission of the instrumental avoidance response led to fewer emitted hurdle-crossings during the regular extinction trials of extinction phase I and re-extinction phase II as compared with the Barrier Group and Control Group. More rapid and stable extinction occurs when the organism is free to emit the instrumental avoidance responses without reinforcement.

REFERENCES

1. ESTES, W. K. An experimental study of punishment. *Psychol. Monog.*, 1944, **57**, No. 31.
2. ESTES, W. K., & SKINNER, B. F. Some quantitative properties of anxiety. *J. Exp. Psychol.*, 1941, **29**, 390-400.
3. MASSERMAN, J. H. Behavior and Neuroses. Chicago: Univ. Chicago Press, 1943.
4. MILLER, N. E. Studies of fear as an acquirable drive: I. Fear as motivation and fear-reduction as reinforcement in the learning of new responses. *J. Exp. Psychol.*, 1948, **38**, 89-101.
5. MOWRER, O. H., & LAMOREAUX, R. R. Avoidance conditioning and signal duration—a study of secondary motivation and reward. *Psychol. Monog.*, 1942, **54**, No. 5.
6. PAGE, H. A., & HALL, J. F. Experimental extinction as a function of the prevention of a response. *J. Comp. Psychol.*, 1953, **46**, 33-34.
7. SCHOENFELD, W. N. An experimental approach to anxiety, escape, and avoidance behavior. *In* P. H. Hoch & J. Zubin (*Eds.*), *Anxiety*. New York: Grune & Stratton, 1950.
8. SIDMAN, M. On the persistence of avoidance behavior. *J. Abn. & Soc. Psychol.*, 1955, **50**, 217-220.
9. SOLOMON, R. L., & BRUSH, E. S. Experimentally derived conceptions of anxiety and aversion. *In* M. R. Jones (*Ed.*), *Nebraska Symposium on Motivation.* Lincoln: Univ. Nebraska, 1956.

Flight Propulsion Laboratory Department
Building 100, Ext. 1730
General Electric Company
Cincinnati 15, Ohio

Published as a separate and in *The Journal of Psychology*, 1959, **47**, 247-258.

AN EXPERIMENTAL STUDY OF THE EFFECT OF HEAD BALANCE ON PATTERNS OF POSTURE AND MOVEMENT IN MAN*

Institute for Applied Experimental Psychology, Tufts University

Frank Pierce Jones, Florence E. Gray, John A. Hanson, and D. N. O'Connell[1]

A. Problem

Jones (5) and Jones and Narva (6) have described an empirical technique for changing the distribution of postural tonus by changing the poise (or tonic balance) of the head. The subject is trained in kinesthetic perception to recognize a hitherto undifferentiated pattern of tension which develops in response to a stimulus to move. The pattern begins with an increase of tension in the muscles of the neck sufficient to exert a slight but appreciable torque upon the head. This torque precedes and accompanies many habitual movements. If the increase in neck-muscle tension is inhibited, the torque on the head can be eliminated from the pattern of response. Such an inhibition results in a redistribution of tonus throughout the trunk and limbs. It is reported by the subject as a reduction in the feeling of weight and a reduction in the effort needed to move or to maintain a position against gravity. It has been conjectured that the mechanism by which this change in tonus is effected is the head-neck reflexes of Magnus.

For experimental purposes the change in tonus can be facilitated by manipulation. When the stimulus to move is presented to the subject the experimenter applies a light upward pressure around the occiput sufficient to counterbalance any torque that may develop. This pressure is sustained throughout a movement or series of movements. Care must be taken to produce only enough pressure to offset the habitual torque, which may vary from person to person and from time to time. If the procedure is successfully carried out, the redistribution of postural tonus and the subjective effect of lightness will be strongly facilitated.

The manipulative procedure is illustrated in Figure 1. In *A* the subject

---

*Received in the Editorial Office on January 12, 1959, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This investigation was supported by a grant from the *USPHS*, Division of Research Grants (*RG*-4836). Publication costs were defrayed by the Tufts Fund for Research in Kinesthesis.

is sitting in his habitual "relaxed" posture. In *B* he has straightened up into his "best sitting" posture. In *C* the experimenter has countered the weight of the subject's head and kept the neck muscles from shortening during the movement. Relative to either of his habitual postures the new, experimental posture may be described as "forward and up." (The three pictures were traced from photographs of a naïve subject taken during a preliminary training period of about 15 minutes.)

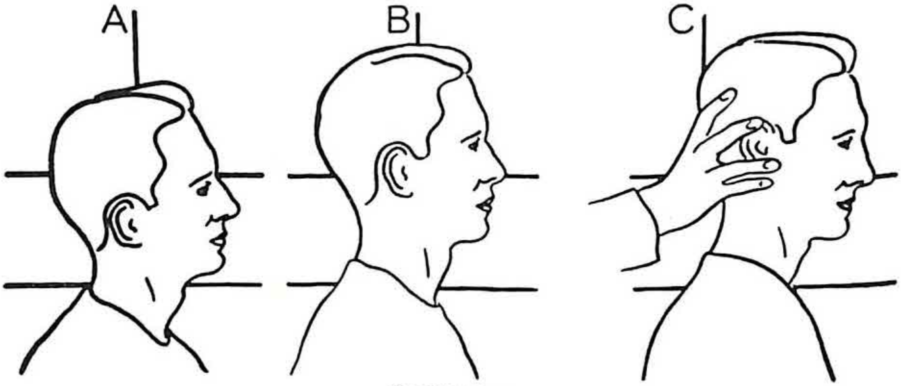


FIGURE 1

Three sitting postures (traced from photographs). *A*, "Sitting comfortably relaxed" (habitual). *B*, "Best sitting posture" (habitual). *C*, "Experimental posture" (into which the subject has been guided by the experimenter).

Once established, the change in tonus tends to persist and to exert an indirect effect upon other patterns of behavior. These behavioral effects have been described elsewhere, and attempts have been made to establish their significance for education and therapy (1, 2, 3). It is difficult, however, to study an indirect effect objectively or to determine with certainty the factors responsible for the change. We have confined our attention here to a study of the change in postural pattern which follows a change in the distribution of tonus. Our principal aims have been: (*a*) to obtain objective evidence that the phenomenon reported by subjects (the kinesthetic effect of lightness and ease of movement) is not due to some form of suggestion but corresponds to a regular, recordable event; (*b*) to find out whether habitual patterns of movement and posture can be modified by training in kinesthetic perception; (*c*) to throw light on the postural mechanisms by which the change is brought about.

B. Experimental Program

The subjects were six male undergraduates from two classes in Voice and Diction. They were paid volunteers ranging in height from 5′4″ to 6′4″, in age from 20 to 26. None had any previous acquaintance with the phenomenon under study.

At the first meeting the subjects were interviewed and the experiment was explained to them as a study of the relation between posture and movement with particular attention to Speech. Nothing was said about the kinesthetic experiences which might be involved. At the second meeting a series of photographs were taken before and after a postural change had been made by the method described above. Single-image photography was used to record static posture and color-coded multiple-image photography (7) to record the postural patterns of movement. Sitting-to-standing was the movement chosen. Though it is a common, everyday movement and can be carried out with the minimum of attention, it presents a "total pattern" which in some degree involves the whole body. For this reason it is well suited for the present study.

Subjective reports were taken at the end of the first photographic session. Appointments were then made with each subject for 20 biweekly meetings of about 30 minutes. At these meetings the subjects were trained in kinesthetic perception to discriminate among the tensional patterns which precede an overt movement and to eliminate those patterns that tend to disturb the poise of the head. From the start, the practical application of the method was stressed, and the subjects were encouraged to apply what they learned to such ordinary activities as walking, talking, studying, or driving a car. Throughout the training period they made oral and written reports of their experiences. At the end, a second series of single-image and multiple-image photographs were taken under the same external conditions as before. All of the photographs were then subjected to statistical analysis.

C. PHOTOGRAPHIC PROCEDURES

The photographs were taken and analyzed by the method described by Jones and O'Connell (8). Markers of reflecting tape indicated the relations between the subject's head and trunk. The points and angles used in the analysis are shown in Figure 2. They are:

A—the seventh cervical vertebra.

B—the mid-point of the Frankfort Plane (drawn from the tragion of the ear to the lowest point of the orbit).

C—the sternal notch.

Epsilon—the angle between the horizontal and the Frankfort Plane.

Theta—the angle between the horizontal and the side *AB* of the triangle.

Eta—the angle between the horizontal and the side *AC* of the triangle (the "chest angle").

Delta—the angle between the Frankfort Plane and the side *AB* of the triangle (the "head-neck angle").

For movement photographs, an additional marker was attached to the thigh.

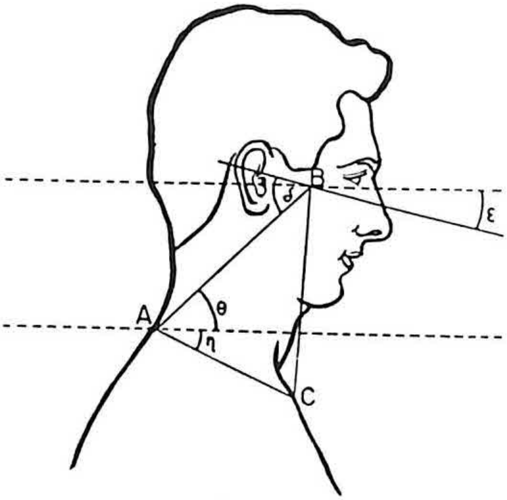


FIGURE 2
Points and Angles Used in Analysis of Movement Patterns

1. *Single-Image Photographs*

For recording "static posture" (a posture assumed in response to a specific verbal stimulus) the subject sat or stood in front of a black screen against which rectangular coördinates had been constructed in white string. The room was darkened and the shutter of the camera opened. The subject's instructions were: (*a*) Sit in your "best sitting posture." (*b*) Stand in your "best standing posture." (*c*) Breathe in deeply and hold it. (*d*) Breathe out deeply and hold it.

Each picture was taken with a single flash from a General Radio Strobolume.

2. *Multiple-Image Photographs*

The multiple-image photographs were taken under similar conditions. While the subject moved from sitting to standing the pattern of his movement was recorded on color film by the Strobolume flashing at 10 *fps* and coded by a color-wheel revolving in synchrony.

For each subject four "habitual" movements were recorded and four "reflex facilitated." For the former, the subject was asked to get up in his ordinary way. Of the latter, two were "adjusted," i.e., the subject got up voluntarily after the experimenter had changed his postural set by the method described above. The other two were "guided," i.e., the postural set was altered as before, and the subject's attention was called to the change in head balance. His instructions were: "Leave your head as it is. Don't make an effort to move. I'll start the movement for you." Standing behind the subject the experimenter started the movement with an upward pressure at the subject's waist. This pressure was sustained until the subject was on his feet. (It should be noted that the experimenter does not make an effort to lift the subject. His function is only to start the movement and to make it easier for the subject to maintain the changed relationship between head and trunk. This can also be accomplished by keeping a hand at the subject's head, as in Figure 1 C.) Of the eight movements, the first two were always "habitual"; the last six were randomized.

D. RESULTS

1. *Subjective Reports*

Five of the subjects reported the kinesthetic effect of lightness after the first session; the sixth (E. H.) did not observe it until after the second. In the course of the training, all of the subjects reported an increase in this effect and a decrease in their feelings of tension and fatigue. They also reported that they had established a more efficient control of themselves in such activities as speaking, studying, and sports, attributing the change to their increased knowledge of postural responses.

Four of the subjects were seniors at the time of the experiment, and no reports were obtained from them after its conclusion. The other two were interviewed a year later. Both reported that they had made continuing use of the postural control which they had established and that its effect in reducing habitual tension and fatigue had persisted.

2. *Single-Image Photographs*

The two sets of single-image photographs taken before and after training were projected and measurements taken of the points and angles shown on the diagram in Figure 2. Significant differences were found in the photographs of sitting posture and of breathing. Differences between the two sets of "best standing postures" were not statistically significant for any of the measures which we used.

a. Best sitting posture. All six of the subjects showed an increase in sitting height after training. As measured from the seventh cervical vertebra, the average increase was 3.5 cm. Increases at the sternal notch and center of the Frankfort Plane were slightly greater. (For some of the subjects it was more difficult to locate the seventh cervical after the training period, since it seemed less prominent than it had been before. There was no difficulty, however, in determining the other two locations.)

b. Habitual sitting posture. A somewhat greater increase in sitting height was found in the "habitual sitting posture," the posture assumed by the subject while he waited for the signal to get up. The measurements, which were taken from the enlarged prints of the movement photographs described below, showed an average increase in sitting height at the 7th cervical vertebra of 4.8 cm.

c. Respiratory postures. The two sets of pictures were compared by measuring the differences in each set between the postures of inspiration and expiration. In the first set (before training), it was found that from inspiration to expiration there was a mean vertical displacement of the head (at the center of the Frankfort Plane) of 5.73 cm. At the same time, angle eta, the angle between the chest and the horizon, increased by 18.9 degrees. In the second set of pictures the vertical displacement of the head between inspiration and expiration was reduced to 0.15 cm, and the increase in the chest angle to 5.9 degrees. Differences between values before and after training were significant at the 0.001 level. The change, which was accompanied by subjective reports of a change in the manner of breathing, points to the need for further investigation. It is included in this report as an additional index of the change in habitual posture.

3. *Multiple-Image Photographs*

Statistical analysis of the multiple-image photographs showed (*a*) a significant difference in the patterns of movement before and after the experimental change in postural set; (*b*) a significant difference in the habitual patterns of movement before and after training.

In the two series, 92 pictures were obtained. (Two were lost in each series because of double exposure.) Enlarged prints were made from the transparencies. By means of the color-coding, the flashes were numbered and lines drawn in pencil between synchronous points on the head and trunk. The changing relations between parts then appeared on the prints as a series of triangles with a cross bar at the apex, the stick marker on the thigh producing a fanlike pattern which began when the subject left the chair. When the patterns had been constructed, it was found that they could be divided by inspection into three groups, corresponding to the three types of movement, Habitual, Adjusted, and Guided. In both series, the Guided patterns could be clearly distinguished from the other two. The Adjusted patterns, though there was considerable variation among them, fell between the Habitual and Guided in Series I. In Series II a change had taken place in the Habitual patterns so that they could no longer be distinguished from the Adjusted. Examples of the three types of pattern are given in Figure 3. All were taken from a single subject (E. C.); and for each type, the first pattern in the series was selected.

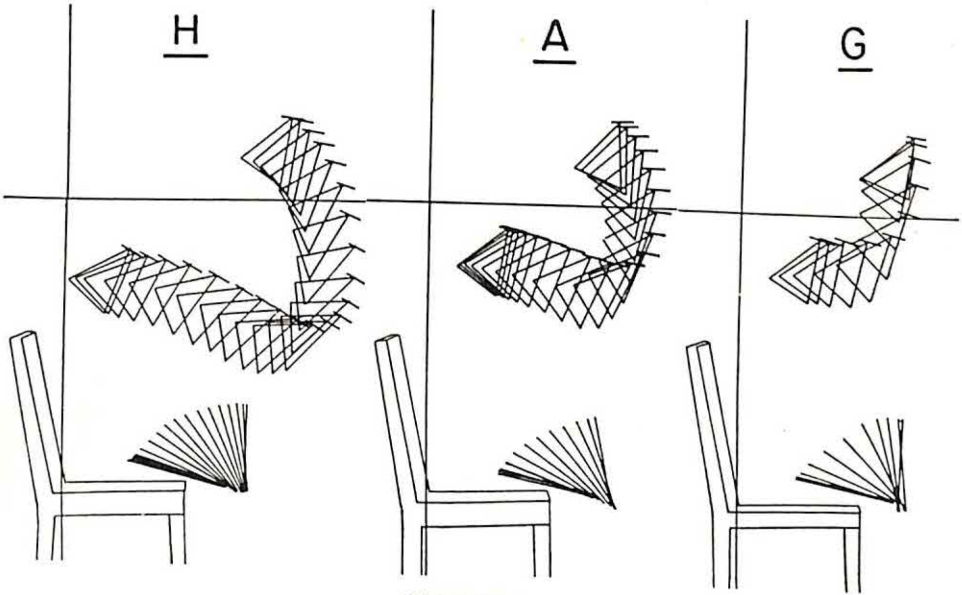


FIGURE 3
"Sit-to-stand." Three movement patterns compared. *H*, "Habitual." *A*, "Adjusted." *G*, "Guided."

In analyzing the patterns the same points and angles were used as in the single-image photographs, plus a fourth point (*D*), the top of the thigh marker. Linear and angular displacements were measured directly from the prints; velocities and accelerations were calculated from the displacements. By plotting the data, all of these relations can be expressed graphically as functions of time or as functions of each other. A partial analysis of the three movements, showing the changes in the head-neck relation is given in Figure 4. The upper pairs of curves show where the head is at any given instant; the lower pairs give its angular relation to the horizon (epsilon) and to the 7th cervical vertebra (theta).

These and other functions were plotted for all of the patterns. To treat the data statistically, we selected such measures as the maxima and minima of various functions and used Mosteller's method of allowances to make a

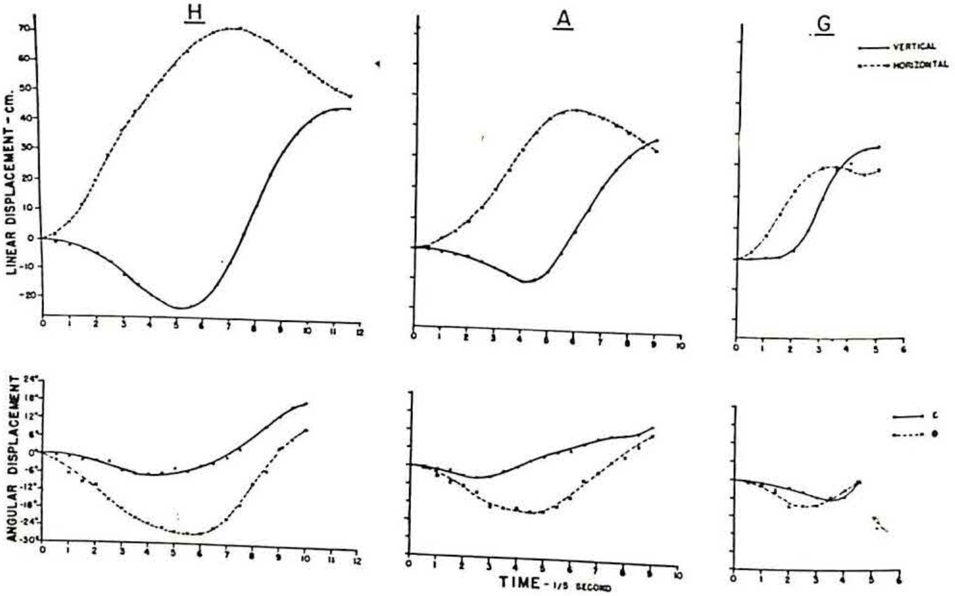


FIGURE 4
GRAPHICAL ANALYSIS OF MOVEMENT PATTERNS IN FIGURE 3 SHOWING LINEAR AND ANGULAR DISPLACEMENTS OF HEAD FROM STARTING POSITION

preliminary survey (9). Among the indices which were found to distinguish the Habitual from the Guided movements, the following appear to throw the most light on the phenomenon itself. In listing them the direction of change is given from Habitual to Guided.

I. Total elapsed time of movement (measured from a cut-off of 99 per cent)—*decreased.*

II. Maximum horizontal velocity of head—*decreased.*

III. Vertical acceleration of head (from point of lowest displacement to point of maximum vertical velocity)—*decreased.*

IV. Lowest vertical displacement of head—*increased* (i.e., became higher).

V. Distance from sternal notch to thigh at moment of rise—*increased.*

VI. Angle delta (the head-neck angle), minimum value—*increased.*

VII. Angle eta (the chest angle), maximum value below the horizontal—*decreased.*

In all of the indices, the Habitual Movements before Training are sharply distinguished from both sets of Guided Movements. They differ from the Adjusted and the Habitual after Training with varying degrees of significance. Bar graphs of the means are shown in Figure 5. *T*-tests were made of the differences. The *t* scores and levels of confidence are given in Table 1.

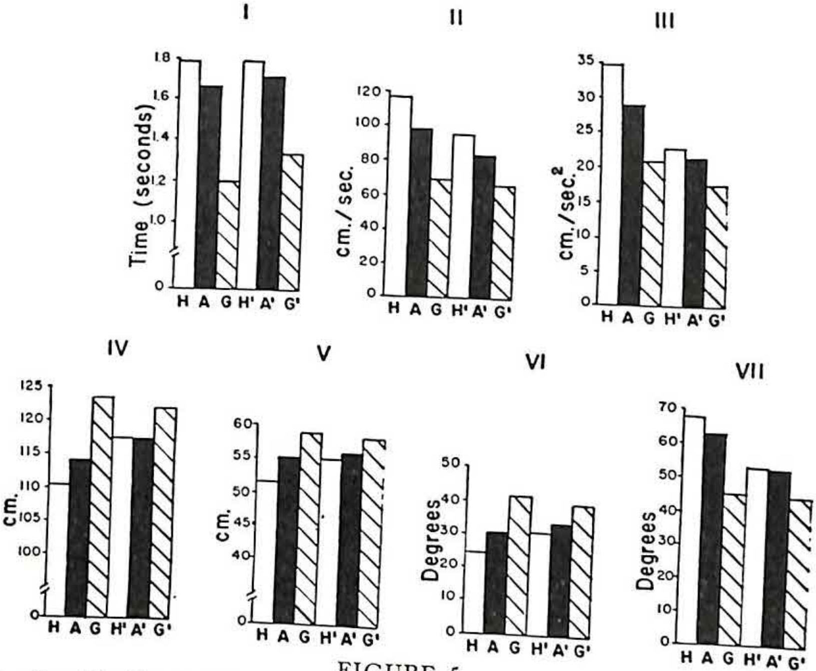


FIGURE 5

Means for "Habitual," "Adjusted," and "Guided" movements before (*H, A, G*) and after (*H', A', G'*) training. I, Total time of movement. II, Maximum horizontal velocity of head. III, Vertical acceleration of head. IV, Minimum distance of head from floor. V, Distance from thigh to sternal notch at time of rise. VI, Angle delta (head-neck angle)—minimum value. VII, Angle eta (chest angle)—maximum value below horizontal.

TABLE 1

DIFFERENCES BETWEEN MEANS IN FIGURE 5. *t* SCORES AND SIGNIFICANCE LEVELS

	Index	H vs A t	H vs G t	H vs H' t	H vs A' t	H vs G' t
I	Total time	0.37	4.91**	0.38	0.17	2.86*
II	Horizontal velocity	3.39*	5.65**	5.90**	11.60**	7.57**
III	Vertical acceleration	2.26	7.07**	4.35**	3.31*	5.64**
IV	Distance of head from floor	1.22	4.76**	1.93	1.75	6.89**
V	Distance from thigh to sternal notch	3.82*	4.61**	2.97*	4.72**	5.04**
VI	Angle δ	2.62*	10.32**	3.04*	4.75**	5.37**
VII	Angle η	1.35	5.74**	3.38*	2.18	3.04*

H = Habitual
A = Adjusted
G = Guided
' = After training
* = Significant at .05 or better.
** = Significant at .01 or better.

E. DISCUSSION

In addition to showing that the two sets of patterns are taken from different movement populations, the statistical analysis provides some of the data that are needed for an operational account of the kinesthetic effect of lightness and ease of movement. This effect, which is perceived most sharply when the movement is elicited as a reflex without the voluntary participation of the subject, is a unity—a Gestalt—and seems difficult at first to describe. In retrospect, however, a certain amount of analysis is possible: (*a*) During the movement, familiar objects are viewed from an unfamiliar angle and the perception of space is altered. (*b*) The rate at which the body and its parts are moving seems to have changed, and there is a different perception of time. (*c*) The movement "feels easier" than the habitual movement. The muscles seem to contract as they do in the knee-jerk or "post-contraction phenomenon," without the feeling of effort which a voluntary movement ordinarily calls forth.

Analysis of the photographic patterns throws light on each of these aspects of the kinesthetic phenomenon:

1. *Orientation in Space*

The patterns show that in the Guided movement head and trunk follow an entirely different trajectory from that of the Habitual. The head remains higher throughout the movement, and its relation to the trunk (as indicated by the measured angles) changes comparatively little, both the backward rotation and the forward thrust being greatly reduced. The trunk also remains at a higher level, and the angle made with the horizon (angle eta) is smaller. If the body is viewed as an energy system, it might be said that its potential energy has been increased by the preliminary change in postural set, and that this energy is conserved to a greater degree in the Guided than in the Habitual movement.

2. *Orientation in Time*

The Guided movements are completed in much less time than the Habitual—sometimes in less than half. This reduction in time is characteristic of a reflex—(it is significant that the Adjusted movements do not show it). Paradoxically, however, there is no corresponding increase in maximum velocity, as there is when the subject is asked to move more quickly. In fact, the horizontal velocity and vertical acceleration of the head actually decrease to a significant degree. This paradox of seeming to move more slowly and yet of reaching the goal more quickly may explain the strangeness in the perception of time which subjects sometimes report. (It is significant that a slower movement does not ordinarily seem easier. All of our subjects, when they were asked to get up more slowly said that it took more rather than less effort to do so. The Guided movements, on the other hand, are described as both "slower" and "easier.")

If the instantaneous velocities of the head are plotted it can be seen that in the Habitual movement, kinetic energy increases as the body moves forward, but that a large proportion of this energy is lost at the moment of leaving the chair. In a successfully guided movement, the pattern is quite different. Velocity curves for an Habitual and a Guided movement are given in Figure 6. In the Habitual a maximum is reached in ½ second, coinciding with the beginning of movement in the thighs. In the next 1/5 second a large part of this velocity is lost, and a minimum is established at the time the head begins to move up. Velocity then increases again until a second maximum is reached. The two-peaked velocity curve is typical of the Habitual movements, occurring in every one of our patterns. Comparison with the curve for a guided movement helps to explain why subjects describe one movement as "jerky" and the other as "smooth." It should be noted that in the Guided pattern the upward movement begins simultaneously in head and thigh.

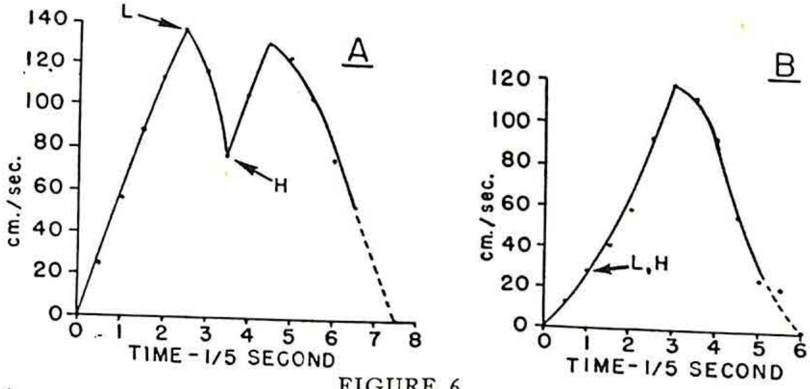


FIGURE 6

Velocity curves for "Habitual" movement (*A*), and "Guided" movement (*B*). *L*, Beginning of upward movement of leg. *H*, Beginning of upward movement of head.

3. *The Feeling of Effort*

The difference in time-space pattern may be enough to explain the subjective feeling of decreased effort. Study of the photographs, however, suggests that another important factor is involved—the length of the muscles which do the work. It is well established that a slight stretch applied to a muscle when it is actively contracting greatly increases its efficiency (4). Our measurements show that for at least two important muscle groups the length is significantly greater during the Guided and Adjusted movements than it is during the Habitual. (It is assumed that the length of the rectus abdominis is directly related to the distance between the upper part of the thigh and the sternal notch; and that the upper trapezius is directly related to the size of the angle delta, since an increase in that angle indicates that the occiput has moved further away from the seventh cervical vertebra.) If similar relations hold for other skeletal muscles, the decrease in the feeling of effort would be adequately accounted for.

In this discussion, we have tried to show that the objective evidence of the photographic patterns complements the introspective reports of the subjects. Both kinds of evidence support the hypothesis that the operations performed before the movement began, whether by the experimenter in changing the relation between head and trunk or by the subject in inhibiting an increase in neck-muscle tension, served to facilitate the postural reflexes.

F. SUMMARY

The effect of head balance on patterns of movement and posture has been studied experimentally in six normal human adults. The subjects were trained in kinesthetic perception to recognize and inhibit tensional responses that disturb the balance of the head. Patterns of posture and movement before and after a change in head-balance were recorded by single-image and multiple-image photography. When the data were analyzed statistically, significant differences were found which correspond to kinesthetic differences reported by the subjects. The findings are discussed in light of the hypothesis that the experimental procedures facilitated the anti-gravity reflexes and changed the total pattern of movement.

REFERENCES

1. ALEXANDER, F. M. The Use of the Self (with an introduction by John Dewey). New York: Dutton, 1932.
2. BARLOW, W. Postural homeostasis. *Ann. Phys. Med.*, July, 1952.
3. COGHILL, G. E. The educational methods of F. Matthias Alexander. *In* Alexander, F. M., *The Universal Constant in Living*. New York: Dutton, 1941.
4. FULTON, J. F. (*Ed.*). A Textbook of Physiology (16th ed.). Philadelphia: Saunders, 1949.
5. JONES, F. P. A mechanism for change. *In* Sorokin, P. (*Ed.*), *Forms and Techniques of Altruistic and Spiritual Growth*. Boston: Beacon Press, 1954.
6. JONES, F. P., & NARVA, M. Interrupted light photography to record the effect of changes in the poise of the head . . . *J. of Psychol.*, 1955, **40**, 125-131.
7. JONES, F. P., & O'CONNELL, D. N. Color coding of multiple-image photographs. *Science*, May 9, 1958.
8. ———. Posture as a function of time. *J. of Psychol.*, 1958, **46**, 287-294.
9. MOSTELLER, F. *In* Lindzey, G. (*Ed.*), *Handbook of Social Psychology*. Cambridge, Mass. Addison-Wesley, 1954. (Vol. I, pp. 304-307.)

Institute for Applied Experimental Psychology
Tufts University
Medford 55, Massachusetts

Published as a separate and in *The Journal of Psychology*, 1959, **47**, 259-266.

PERCEPTION OF VERTICALITY IN MIDDLE AND OLD AGE*[1]

Department of Psychology, Clark University

PETER E. COMALLI, JR., SEYMOUR WAPNER, AND HEINZ WERNER

A. INTRODUCTION AND PROBLEM

Recently we published a monograph (2) devoted to the problem of perceptual development, or more specifically to the changes in spatial organization that occur during the course of growth. The study was undertaken within sensory-tonic theory which deals with perception in terms of "object: organism" relationships and general developmental theory that specifies the ontogenetic changes in this relationship in terms of differentiation.

A number of experimental situations were employed in that study of perceptual development, one of which, pertinent to the present paper, concerns the perception of verticality. With subjects ranging in age from 6 to 20 years, two significant developmental trends in the perception of verticality were found. One concerns the effect of body tilt: within the younger age levels, a luminescent rod in a darkroom is perceived as vertical when tilted to the side at which the body is tilted, in contrast to older age levels where it is perceived as vertical when tilted opposite the side of body tilt. The second effect is that of "starting position," i.e., the rotation of the apparent vertical to the side at which the rod is placed at the beginning of a trial; this effect is greatest for the youngest age levels and decreases with increase in age. Without going into detail, the results were interpreted as being consonant with the general developmental principle that with increase in age there is increasing differentiation of self and object. Specifically, it was assumed that this lack of differentiation at the early age levels is manifested in two ways: (*a*) by "egocentricity," i.e., determination of the object world through the self (body) as referent, as evidenced by the effect of body tilt; and (*b*) by "stimulus boundedness," i.e., inordinate impact of initially presented stimuli, as evidenced by the effect of starting position.

The present study again is concerned with developmental changes in spatial organization, as reflected in perception of verticality, but deals with such changes that occur beyond the 20 year age level and in particular in old age; its further purpose is to compare these changes with those obtaining at younger age levels so that a total picture will be gained of the development of spatial organization between 6 to 80 years of age.

---

*Received in the Editorial Office on January 25, 1959, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This investigation was supported by a research grant M-348 from the National Institute of Mental Health, of the National Institutes of Health, Public Health Service.

1. *Procedure*

The task for *S* was to adjust, in a dark room, a luminescent rod so that it appeared vertical. The rod, 30 in. long and 1 in. wide, was rotated in the fronto-parallel plane and pivoted at its center; its angular displacement from the perpendicular was measured in half degrees. The task of adjusting the rod to apparent verticality was carried out under 12 conditions, viz., three conditions of body position (30° left tilt; erect; 30° right tilt) in combination with four conditions of starting position of the rod (30° left; 10° left; 10° right; 30° right). A different random order of presentation of trials was used for each *S*.

S sat in a chair which was pivoted below the seat permitting left and right tilts. The chair was provided with head rest, sides and back, to make certain that the body was held firmly in place. While *S* was blindfolded, the chair and rod were set into positions required for the particular trial. *S*, after opening his eyes, instructed *E* to move the rod till it appeared vertical to *S*. *E* followed *S*'s instructions by moving the rod first in 2° steps, and finally in smaller steps. After adjustment of the rod to a physical position in which it appeared vertical, *S* was again blindfolded until the beginning of the next trial.

2. *Measures Employed*

The position in which the rod looked vertical to *S* (position of apparent vertical) was measured in degrees of deviation from the objective vertical (plumb line). Arbitrarily, an angular position clockwise of objective vertical, as viewed by *S*, is designated "+" (right tilt); and an angular position counter-clockwise of objective vertical as "—" (left tilt).

3. *Subjects*

A total of 75 men between the ages 20 and 80 were tested. The age distribution is shown in Table 1.

TABLE 1

	Age						
	20:0 24:11	25:0 29:11	30:0 34:11	35:0 39:11	40:0 44:11	45:0 49:11	65:0 80:0
Number of subjects	10	10	10	10	10	10	15

The *S*s between the ages 20 and 50 were drawn from the Clark University Evening College and the local community; those between 65 and 80 years were drawn from a community old-age club.

B. RESULTS

1. *Effect of Body Tilt*

The over-all results are presented in the graph shown in Figure 1. The graph plots the mean positions of apparent vertical under the three body positions for the various age levels. Considering the effect of body tilt for ages 20-50 there is confirmation of a finding that has been substantiated in many studies, viz., with body tilted left the apparent vertical is rotated to the right, and with the body right the apparent vertical is rotated to the left (4). This effect of body tilt changes markedly with the oldest group (65-80 years): with body tilt left, the physical position of the apparent vertical is rotated to the left; and with body tilt right, it is rotated to the right. In other words, a reversal occurs in the tilt relation of body and rod at old age: for the 20-50 year groups the position of apparent vertical is tilted relatively to the *opposite* side of body tilt; in contrast, for the 65-80 year group, the position of apparent vertical is tilted relatively to the *same* side as body tilt. This developmental shift is statistically significant. As shown in Tables 2 and 3 the interaction $A \times B$ (age by body tilt) is significant beyond the .01 level of confidence.

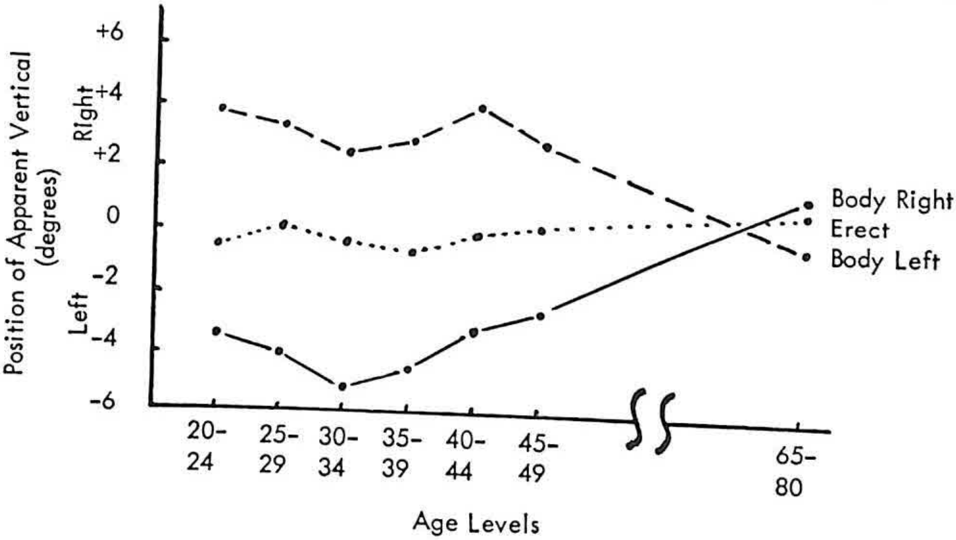


FIGURE 1
DEVELOPMENTAL CHANGES IN EFFECT OF BODY TILT ON PERCEPTION OF VERTICALITY BETWEEN 20 AND 80 YEARS OF AGE

TABLE 2

DEVELOPMENT AND PERCEPTION OF VERTICALITY

Source of variation	F-tests df		Mean square	F	P
Between Subjects (S)	74		32.05	—	—
Between Age Levels (A)		6	43.48	1.40[a]	> .05
Between S's within A's		68	31.04	—	—
Within Subjects	825		46.39	—	—
Between Body Tilt (B)		2	2090.97	34.65[b]	< .01
Between Starting Position (St)		3	3273.43	83.84[c]	< .01
B × St		6	278.64	29.96[d]	< .01
A × B		12	162.18	2.68[b]	< .01
A × St		18	20.83	0.53[c]	> .05
A × B × St		36	8.59	0.92[d]	> .05
Pooled S's × B		136	60.34	—	—
Pooled S's × St		204	39.04	—	—
Pooled S's × (B × St)		408	9.30	—	—
Total	899				

[a] Tested against "Between S's within A's."
[b] Tested against "Pooled S's × B."
[c] Tested against "Pooled S's × St."
[d] Tested against "Pooled S's × (B × St)."

TABLE 3

DEVELOPMENT AND PERCEPTION OF VERTICALITY

A × B F-test: P < .01		Means (degrees)* 30° Left	Body tilt (B) Erect	30° Right
Age (A)	20-24	+3.8	—0.6	—3.4
	25-29	+3.3	0.0	—4.0
	30-34	+2.4	—0.5	—5.1
	35-39	+2.9	—0.8	—4.5
	40-44	+4.0	—0.2	—3.2
	45-49	+2.8	+0.1	—2.6
	65-80	—0.4	+0.8	+1.3

A × St F-test: P > .05		*Starting Position (St)* 30° Left	10° Left	10° Right	30° Right
Age (A)	20-24	—4.5	—2.6	+2.2	+4.7
	25-29	—3.5	—1.9	+2.4	+2.1
	30-34	—5.0	—2.8	+1.7	+1.8
	35-39	—4.7	—3.2	+1.5	+3.2
	40-44	—5.2	—2.3	+3.1	+5.3
	45-49	—3.9	—3.4	+2.9	+4.7
	65-80	—3.0	—2.0	+3.4	+3.8

*A plus sign indicates a position of the apparent vertical clockwise of the plumb line; a minus sign indicates a position counterclockwise.

2. *Starting Position*

As mentioned above, by starting position effect is meant the influence of the position of the rod at the start of a trial on the final position in which it is seen as vertical. In Figure 2, the mean position of apparent vertical under the four starting positions of the rod is shown for the various age levels. The graph shows the effect demonstrated in many other studies with adults, viz., the physical position of apparent vertical is rotated relatively to the left under left starting position and relatively to the right under right starting position (3). Though there are fluctuations in the curves, no consistent developmental change is apparent. In this connection it should be noted that the interaction $A \times S$ (age by starting position) is not significant (see Tables 2-3).

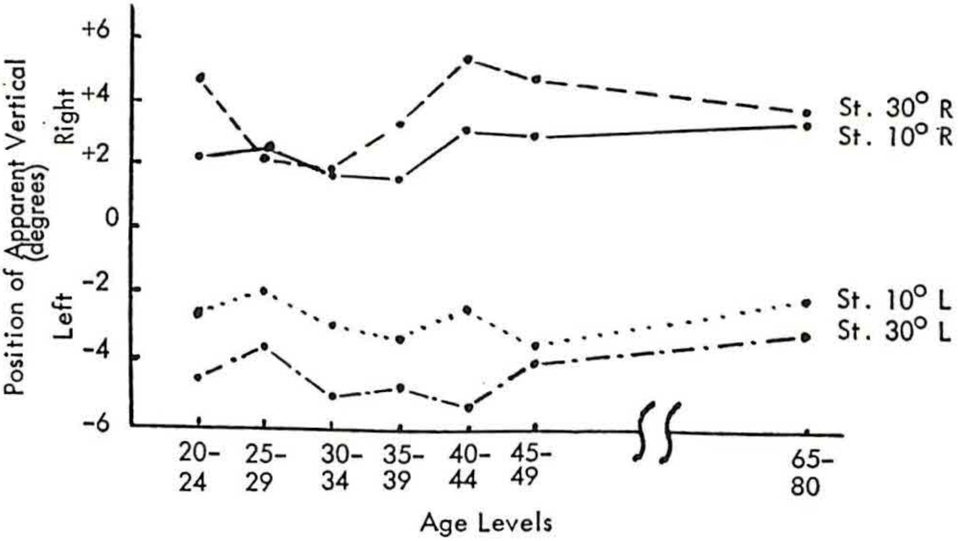


FIGURE 2
DEVELOPMENTAL CHANGES IN EFFECT OF STARTING POSITION ON PERCEPTION OF VERTICALITY BETWEEN 20 AND 80 YEARS OF AGE

3. *Comparison with Developmental Changes at Age Levels 6-20*

If one combines the data of the present study with those obtained previously for younger ages a very interesting result becomes apparent (see Figure 3). For both young boys, 6-15, and older men, 65-80, the apparent vertical is located to the same side as body tilt; in contrast, the age groups between 16 and 50 show the opposite effects of body tilt, viz., the apparent vertical is located to the opposite side of body tilt. In sum, at age levels 6-15, the

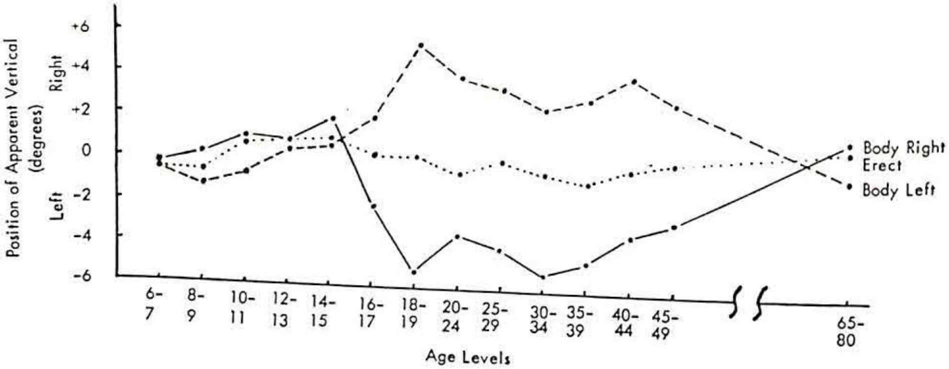


FIGURE 3
Developmental Changes in Effect of Body Tilt on Perception of Verticality Between 6 and 80 Years of Age

apparent vertical is located to the *same* side as body tilt. Beyond the age of 15 the apparent vertical is located *opposite* body tilt. This relationship begins to reverse itself at about 45 years of age so that for the 65-80 year group the same relationship found at the early age levels obtains.

The developmental results for starting position from ages 6 through 80 are presented in Figure 4. The curves clearly indicate a significant decrease in the starting position effect between 6 and 19 years of age. Though there are fluctuations in the curves after 19 years of age, no significant trend is apparent.

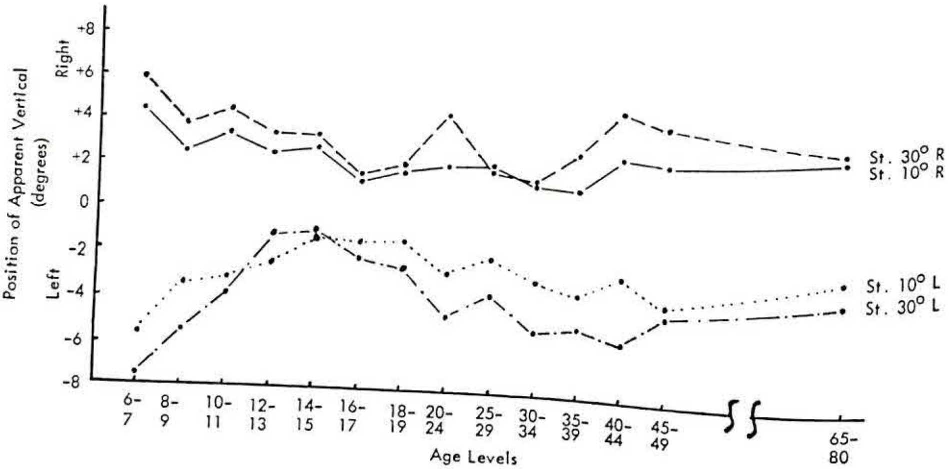


FIGURE 4
Developmental Changes in Effect of Starting Position on Perception of Verticality Between 6 and 80 Years of Age

C. SUMMARY AND DISCUSSION

The present investigation, together with the previous study concerned with early age levels, contributes to our knowledge concerning changes in spatial organization that occur through life. Striking differences in effect of body tilt depending on age level were found: for young boys 6-15 the apparent vertical is located to the *same* side as body tilt; between 16 and 50, however, the opposite effect occurs, viz., apparent vertical is located to the *opposite* side of body tilt; finally in older men, 65-80, the apparent vertical is again located to the *same* side of body tilt. The developmental changes in effect of starting position were found to occur only within the younger age range, viz., starting position effect is greatest at the youngest age level and decreases markedly until the 19 year level; following this there are no consistent developmental changes throughout the age levels studied.

Both of these effects occurring during the period of growth through adolescence were interpreted by us as an expression of the orthogenetic principle which states that the development proceeds from a state of globality and lack of differentiation—in this case between self (own body) and object—to a state of increased differentiation. We interpret our results in accordance with this principle by assuming that at early stages of development lack of differentiation of self and object manifests itself in two ways: (*a*) by egocentricity—i.e., the formation of a frame of reference whereby the stimuli are interpreted in terms of body position; and (*b*) by stimulus boundedness—which refers to the inordinate impact of object stimuli.

As to our findings concerning the old age groups it is here suggested that the reversal in the tilt relation of body and rod indicates the reverting to a more egocentric organization of space, i.e., once again the object world is determined through self as referent. It would, however, be erroneous to assume that the older person operates in a manner identical with that of the child. An important difference is that the stimulus boundedness, evidenced by the large starting position effect, characteristic of the child is not present in the older person. That means, in contrast to the child, the older person maintains a relatively stable spatial frame of references not subject to the whim of every change in stimulation.

One might argue that the greater egocentric determination of space in old age is indicative of senility in terms of deterioration. We would, however, caution against such an interpretation in the light of the fact that large starting position effects do not occur. In this regard we should like to draw attention to studies on perception of verticality which suggest that stimulus boundedness, as indicated by large starting position effects, are associated with brain damage (1). Therefore, we would argue that the absence of increased starting position effects in old age coupled with egocentric position of the apparent vertical would define changes in spatial organization occurring in old age as part of normal ontogenetic development. It remains for future studies to determine whether a distinction can be made between normal and pathological perceptual changes in old age, one evidenced by the absence and the other by the presence of large starting position effects.

REFERENCES

1. TEUBER, H. L., & LIEBERT, R. S. Specific and general effects of brain injury in man. *A.M.A. Arch. Neurol. & Psychiat.*, 1958, **80**, 403-407.
2. WAPNER, S., & WERNER, H. Perceptual Development. Worcester: Clark Univ. Press, 1957.
3. WERNER, H., & WAPNER, S. Experiments on sensory-tonic field theory of perception: IV. Effect of initial position of a rod on apparent verticality. *J. Exp. Psychol.*, 1952, **44**, 68-74.
4. WERNER, H., WAPNER, S., & CHANDLER, K. A. Experiments on sensory-tonic field theory of perception: II. Effect of supported and unsupported tilt of the body on the visual perception of verticality. *J. Exp. Psychol.*, 1951, **42**, 346-350.

Department of Psychology
Clark University
Worcester 10, Massachusetts

Published as a separate and in *The Journal of Psychology*, 1959, **47**, 267-276.

AN INVESTIGATION OF THE EFFECTS OF REWARD AND PUNISHMENT ON VISUAL PERCEPTION*

Department of Psychology, University of Denver

FRANK S. BEATTY, LAWRENCE E. DAMERON, AND JOEL E. GREENE

A. PROBLEM

This investigation represents another in the growing number of attempts to establish the importance of the individual's experiences with reinforcing stimuli on his subsequent related perceptions. This general area of study (and the theoretical thinking) has been labeled by Allport (1) as "directive-state" and represents the familiar practice of contrasting innate biological and physical determinants of perception with learned and functional influences.

Jackson (2), Rock and Fleck (3), and Schafer and Murphy (4) reported studies in which they investigated the specific rôle of reward and punishment in a visual figure-ground relationship. By assuming that the acquisition of money was rewarding to the *S* and the retraction of money was punishing, they attempted to demonstrate that motivational determinants affect the content of an individual's perception. Their technique was to build up a strong association between certain visual patterns and reward, and between other visual patterns and punishment; later a "reward" pattern was fitted with a "punishment" pattern in such a way that one perception could occur only at the expense of the other. Some of the results of these studies lend support to the hypothesis that rewarded patterns will be perceived significantly more times than punishment patterns, but the data are by no means unequivocal.

The hypothesis tested in the present study was that the perceptual threshold of visually presented rewarded profiles of faces is lower than that of profiles which had been associated with punishment. The perceptual threshold was determined by measuring in foot-candle units the intensity of light on the screen which had been projected through the profile slides.

It is realized that the rewards and punishments used in the laboratory may not be rewarding or punishing to the *S*s studied, though generally speaking they may be so sociologically. Therefore, this experiment was modified to allow for independent consideration of the results found for

---

*Received in the Editorial Office on January 26, 1959, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

those *S*s who report themselves as feeling rewarded or punished, and those who report to the contrary. By so doing, it was felt that a more adequate understanding of reward-punishment influences would be achieved. Perhaps, also, there would be less of the contamination that occurs so frequently in such experiments, namely that an experiment becomes a study of *what* is rewarding and punishing as well as the *effects* of reward and punishment.

The criteria for determining the effects of reward and punishment were different from the criteria for previous studies, necessitating several changes in the basic design of the experiment. First, a pretraining series was incorporated to assure that the profiles and their names had been learned before the posttraining series. Second, four individual profiles were presented separately throughout the study to assure that the material was perceptually equivalent and familiar to the *S*s. Also, it was believed that any interpretative factors which could have affected the results were reduced by presenting the profiles individually rather than in a reversible figure-ground configuration. Third, reward and punishment might have been more apt to be present due to the giving and taking away dollar bills rather than the coins used in previous studies. Fourth, to arrive at some estimate of the validity of the assumption that receiving and returning money was rewarding and punishing to *S*, a brief interview was carried out after the termination of the main experiment.

B. METHOD

1. *Subjects*

The *S*s used in this experiment were 30 undergraduate college students, 16 males and 14 females. They were randomly assigned in equal numbers to two experimental groups. Six *S*s failed to learn to criterion during the pretraining series and were eliminated from the experiment. The results are based upon 24 *S*s, 6 males and 6 females in each group.

2. *Apparatus*

Apparatus consisted of a Keystone overhead projector-type tachistoscope with a 750 w. bulb; a 130 v. variac; a light meter; four 3¼ in. × 4 in. glass slides on which were drawn different vertical, outline facial contour profiles; and a screen onto which the profiles were projected. All profiles faced toward the right and measured 2¼ in. high and 1 in. wide when projected on the screen. Profile *A* was named Bob; *B*, Jim; *C*, Ted; and *D*, John. The criteria for selecting the names were that each name be easy to pronounce and familiar to the *S*s—each being a common, one syllable name. The purpose of these names was to enable the *S*s to identify correctly the profiles later in the experiment. It was felt that these profiles were superior to other types of stimulus material because they were assumed to be perceptually equivalent, easy to learn, and unfamiliar to the *S*s, thus controlling for prior learning. It was determined before the study began that the slides all admitted the same amount of light to the screen.

The tachistoscope was placed on a table, the lens being 8 in. from the screen. The projected image on the center of the screen was about 3 ft. away and approximately eye-level for each *S* when he was seated. Black shades were placed over the windows, and the constancy of room illumination was established by a check with a light meter.

Keeping exposure time constant at .01 sec. the variac was set at 20 v. and increased at the rate of 2 v. per exposure until the profile was identified by the *S*. If the profile was correctly identified the tachistoscope was shut off, the profile removed, another profile placed on the tachistoscope for exposure, and then exposed. If, however, the profile was incorrectly identified, the same procedure was followed except that the profile was presented again. In this way the *S* was never given any indication whether his identifications were correct or incorrect. Thus, any identifications made by a process of elimination were minimized.

3. *Pretraining Series*

Each *S* was seated in a chair at a table and given the following oral instructions:

> This is an experiment in visual perception. You will be shown profiles of four faces on the screen in front of you. As each profile is shown a name will be called out. Please observe each profile and listen to its name carefully so that you will remember it. You will be tested on them later.

Each profile was then projected on the screen for 1 sec. at 240 foot-candles, and the name which *E* had assigned to the profile was called out simultaneously. Each profile was presented in each serial position an equal number of times, and the same order was administered to each *S*. After each profile and name had been presented twice, the following instructions were given orally: "Now, as each profile is shown, I want you to tell me the name of the face. All right, here is the first one." Each profile was then presented until *S* had correctly identified the profile four times. As the profile was learned to criterion, it was eliminated from the series to minimize the effects of over-learning. To insure equal performance, a final series of the four profiles was given. All *S*s identified all profiles correctly on this series.

4. *Training Series*

The purpose of the training series was to present a situation whereby the profiles would be associated with reward or punishment. Each profile was exposed on the screen for 1 sec. and assigned the predetermined reward or punishment; the name being called out simultaneously. The profiles were presented in a systematized random order 12 times making a total of 48 presentations.

Before the training series had begun, a stack of dollar bills was placed on the table with one dollar placed directly in front of *S*. The following instructions were read:

> I am going to expose each of the profiles rapidly a number of times. Before I expose each profile, you will call out a number from one to four. I will say the name of the face which happens to fall on the number you call and expose it on the screen. The faces won't necessarily have the same number each time, because I may switch them around after each presentation. (*E* then demonstrated how he could change each profile to a different number by using numbered papers on which the profiles were placed. Of course, the system was predetermined unknown to *S*.) If that face wins, you may take two dollars from the money on the table; if it loses you give me back one dollar. When either of two of the four profiles is called, you win; when either of the other two is called, you lose. It's like a game of chance. You will be permitted to keep a percentage of your winnings after the experiment is over. Do you understand?

The "gambling" method was employed because it was believed that this type of activity would be interesting and permit *S* to become involved in the situation. To make the money as effective as possible, each *S* was permitted to take or give back the money himself. Each *S* "won" $25 at the end of the series. Twelve *S*s received two dollars each time *A* and *D* were presented and returned one dollar each time *B* and *C* were presented. The procedure was reversed for the other 12 *S*s; *B* and *C* were rewarded, *A* and *D* were punished. Two dollars were given as reward and one dollar taken away as punishment so that the *S*s would always increase their winnings, thus eliminating, as much as possible, discouragement on the part of the *S*s.

5. *Posttraining Series*

This series served to test whether *S*s rewarded for facial profiles by being given two dollars per profile, would perceive these at significantly lower illumination thresholds than *S*s punished for other profiles by having one dollar taken away. The illumination required to perceive each profile was obtained using the Method of Limits to determine the lower absolute threshold. Because of the nature of the experiment, only ascending series were used. It was believed that the use of a descending series would have magnified the error of habituation.

The following instructions were read aloud:

> I am now going to flash the profiles on the screen very rapidly. As soon as you think you recognize the face, call out the name. If you don't recognize it, just say, "No." Do you understand?

The method of systematic randomization for presentation of the profiles was used during this series. Thus 16 threshold values were obtained for each *S*.

6. *Postexperimental Interview*

To arrive at some estimate of the validity of the assumption that receiving and returning money was rewarding and punishing to *S*, a brief interview was conducted after the termination of the main experiment. Though aware of the problems inherent in using subjective report to establish the presence of motivation, it was felt that some insight could be gained in this manner, and that these findings could, in turn, be subjected to empirical testing.

The following questions were asked:

1. What do you think is the purpose of this experiment?
2. How did you feel when you received money for the faces?
3. How did you feel when the money was taken away from you?
4. Is there anything else you would like to mention concerning the experiment?

It is felt that *S*s were given no cues which might have influenced the responses to the questions, inasmuch as *E* did not know the results of the posttraining series until after the entire experiment had been completed and the data analyzed.

C. RESULTS

All threshold values were converted to foot-candle units for purposes of statistical analysis. Inspection of the distributions of the data disclosed that parametric statistical methods would be inappropriate. Therefore, several nonparametric statistical methods were employed. Means, medians, and standard deviations in foot-candle units are presented in Table 1.

Using the Wilcoxin Matched Pairs Signed Ranks Test no significant differences in threshold value were found between all rewarded and all punished profiles. However, this comparison was made under the assumption that all profiles were equivalent. In a pilot study the profiles had approxi-

TABLE 1
DESCRIPTIVE STATISTICS OF THRESHOLD VALUES IN FOOT-CANDLE UNITS FOR GROUPS, CONDITIONS, AND PROFILES

Profile		*A*	*D*	*B*	*C*	*A* + *D*	*B* + *C*
Grp. I (*A* and *D* Rewarded)	*M*	41.90	48.40	40.10	41.50	90.20	81.60
	SD	5.76	11.00	9.86	11.08	16.42	15.52
	Mdn	41.00	44.00	38.00	39.00	85.50	78.00
	Range	33-52	39-74	30-67	29-75	72-126	59-113
Grp. II (*A* and *D* Punished)	*M*	66.30	99.70	62.50	51.30	166.10	113.70
	SD	49.36	26.17	38.31	15.37	130.09	51.88
	Mdn	46.50	72.00	49.00	49.50	118.50	100.00
	Range	34-208	39-338	34-178	34-86	76-546	69-264

mately the same threshold values. Further analysis, however, indicated that the assumption of equality among the profiles was doubtful.

The paired profiles *B* and *C* (paired in terms of both being rewarded and punished for the same *S*s) were perceived at a lower threshold than paired *A* and *D*. This difference was apparently a function of Profile *D* rather than *A* combined with *D*, inasmuch as there were no significant differences between *A*, *B*, and *C*. Differences significant at or beyond the .01 level of confidence were found, however, between *D* and the others, which were all perceived at lower thresholds than *D*. It seems that Profile *D* was different from the other profiles in some unknown respect making a comparison between paired profiles invalid.

Therefore, intra- and inter-profile comparisons were made for *A*, *B*, and *C*. That is, rewarded Profile *A* was compared with punished Profile *A*; each of these were compared with punished and rewarded Profiles *B* and *C*; and *B*'s were compared with *C*'s. Profile *D* was compared with itself since it was shown to be different perceptually from the other profiles. The Mann-Whitney two-tailed test was applied in comparisons involving independent groups. Results of this analysis indicated that *A* punished, *B* rewarded, and *C* rewarded were all perceived at higher thresholds, significant at or beyond the .05 level, than *B* punished and *C* punished. Thus, it seems the effects of reward and punishment on Profiles *B* and *C* were reversed from the predicted direction; reward seemed to elevate the threshold, whereas punishment seemed to lower it.

It appeared to be more meaningful to discover whether reward and punishment were effective by determining whether each *S* perceived a profile for which he had been rewarded at a lower threshold than one for which he had been punished. Therefore, those *S*s who were rewarded for *A* and *D*, punished for *B* and *C*, were compared with each other as a group, and those who perceived the profiles under the opposite reward-punishment condition were compared as a group. This analysis revealed that neither group perceived profiles for which they had been rewarded at significantly lower thresholds than those for which they had been punished or vice versa, except in the case of Profile *D*.

When the groups were compared with each other, however, those *S*s who were rewarded for *A* and *D*, punished for *B* and *C*, perceived all profiles at lower thresholds, than those *S*s who were punished for *A* and *D*, rewarded for *B* and *C*. This difference was significant at the .02 level of confidence, and could not be attributed to differences between the threshold values of males and females nor to differences in learning. In the pretraining series, all profiles required approximately the same number of trials to learn the profile and name to criterion.

The thresholds at which the profiles were perceived were apparently a function of the *S*s rather than a function of the predicted effects of reward and punishment or differences among profiles, except in the case of *D*. That the threshold values were a function of the *S*s seems to be substantiated by an analysis of the postexperimental interview.

D. POSTEXPERIMENTAL INTERVIEW

During the questioning after the main experiment, 10 *S*s expressed that they liked to receive money and disliked losing it. Four others expressed that they liked to receive it and felt neutral about losing it. The remaining 10 expressed no involvement in either receiving or losing the money. Assuming that those expressing involvement had been psychologically rewarded and punished, it would be expected that they would have perceived profiles for which they had been rewarded at significantly lower thresholds than those *S*s who expressed no involvement. Conversely, it would be expected that those who expressed involvement over losing money would have perceived those profiles for which they had been punished, at significantly higher thresholds than those who expressed no involvement. These comparisons revealed no significant differences.

When the data were analyzed in terms of differences in performance between involved and noninvolved *S*s, however, a rather striking finding was obtained. Those *S*s who expressed involvement perceived all of the profiles, whether rewarded or punished, at lower thresholds than those *S*s who expressed no involvement. The difference was significant at the .004 level of confidence. The involved *S*s also displayed less variability (the range between *S*'s highest and lowest thresholds) than the noninvolved *Ss*—this difference being significant at the .01 level of confidence.

Involvement or lack of it was not expressed as a function of the profiles but could not be attributed to chance. Although a majority of those *Ss* who expressed involvement perceived rewarded *A* and *D*, punished *B* and *C*, Fisher's exact probability method for 2 × 2 tables showed no significant difference between the number of involved and noninvolved *Ss* perceiving the opposite reward-punishment conditions.

E. Discussion

The data indicated that there was no significant difference between all profiles which had been rewarded and all which had been punished, attributable to an apparent reversal of the effects of reward and punishment on the profiles. It was found that Profiles *A* and *D* were affected more by reward than punishment in that they were perceived at lower thresholds when rewarded than when punished. The opposite results were found for *B* and *C* which were perceived at lower thresholds when punished than when rewarded.

It is apparent that reward and punishment as hypothesized did not affect the perception of the *Ss* in either group, in view of the fact that all profiles, except *D*, were perceived at approximately the same thresholds within each group. The reversed perceptual thresholds for *B* and *C* were accounted for by the differences found between the two groups of *Ss* rather than by the reversible effects of reward and punishment.

It was found through subjective report that those *Ss* who expressed involvement in receiving or losing money perceived the profiles at significantly lower thresholds and displayed less variability in their thresholds than those who expressed no involvement.

The evidence would lend itself to the conclusion that although money did not serve as a means of reward and punishment in the sense of effecting a bi-directional change in perception, it did influence thresholds. It is possible that both receiving and returning money provided a "set" for the involved *Ss*. This set facilitated the perception of the profiles. It is speculated that the results can be accounted for by assuming that, once involved in the task, the *S* had a set to "do well" and hence improved his performance. The use of money as reward and punishment, as far as the conditions of this experiment are concerned, appears to mediate some kind of set to achieve rather than to serve directly as a reinforcing agent.

F. SUMMARY

It was the purpose of this study to determine whether *S*s who had been rewarded upon the presentation of certain profiles perceived those profiles at significantly lower illumination thresholds than when punished upon the presentation of other profiles. It was the further purpose to determine whether money served as reward and punishment in this experimental situation by obtaining a subjective report as to each *S*'s involvement concerning the acceptance and return of money.

The experiment consisted of presenting four vertical, outline, contour, facial profiles tachistoscopically, to 24 undergraduate college students. The profiles and their names were learned during a pretraining series; associated with reward and punishment during a training series; and presented during a posttraining series to determine the effects of reward and punishment on the *S*'s perception. The experiment was followed by an interview to determine the *S*'s involvement concerning the acceptance and return of money.

No significant differences were found between all profiles which had been rewarded and all which had been punished. Reward seemed to have been more effective for Profiles *A* and *D*; punishment seemed more effective for *B* and *C*. The reversed perceptual thresholds for *B* and *C* were accounted for by the differences found between the two groups of *S*s rather than by the reversible effects of reward and punishment. It was also found that *S*s who perceived the profiles under the same set of conditions did not perceive the profiles for which they had been rewarded at significantly lower thresholds than those for which they had been punished, except in the case of *D*, which was found to be different in some unknown manner, from the other profiles.

A comparison between those *S*s who expressed involvement and those who did not, uncovered the finding that involved *S*s perceived the profiles at significantly lower thresholds, and displayed less variability between thresholds, than noninvolved *S*s. It is speculated that this result can be accounted for by assuming that, once involved in the task, the *S* had a set to "do well" and hence improved his performance. The use of money as reward and punishment, as far as the conditions of this experiment are concerned, appears to mediate some kind of set to achieve rather than to serve directly as a reinforcing agent.

REFERENCES

1. ALLPORT, F. H. Theories of Perception and the Concept of Structure. New York: Wiley, 1955.
2. JACKSON, D. N. A further examination of the rôle of autism in a visual figure-ground relationship. *J. of Psychol.*, 1954, **38**, 339-357.

3. Rock, I., & Fleck, F. A re-examination of the effect of monetary reward and punishment on figure-ground perception. *J. Exp. Psychol.*, 1950, **40**, 766-776.
4. Schafer, R., & Murphy, G. The rôle of autism in a visual figure-ground relationship. *J. Exp. Psychol.*, 1943, **32**, 335-343.

Sonoma State Hospital
Eldridge, California

Published as a separate and in *The Journal of Psychology*, 1959, **47**, 277-288.

THE BEHAVIORAL WORLD*

Department of Neurology and Psychiatry, Northwestern University Medical School

HARVEY NASH[1]

Charlie Chaplin, in one of his classic performances, portrayed a "Gold Rush" prospector who is so crazed by lack of food that his perceptions and actions are transformed to provide satisfaction for his hunger, in fantasy. The prospector boils his shoes and, knife and fork in hand, proceeds to ingest the shoes with gusto, *as if* he were eating a steak. The shoelaces are consumed separately; the prospector manipulates the laces in such a manner that the audience cannot but feel that he believes himself to be eating spaghetti. This situation is depicted in Figure 1, where a distinction is made between the objects and events of the real world, and the objects and events which, while not present in fact, are postulated in order to "make sense out of" the prospector's actions.

Much of the behavior of an individual is ordered around objects which may or may not exist "in fact." The need to establish and to maintain an ordering around focal objects varies with the age of the individual, with his external circumstances, with his character structure, with any psychopathology from which he may be suffering, and so forth. The object-centeredness of the individual's behavior is depicted in Figure 2, which is discussed in the following paragraphs.

A. THE REAL (GEOGRAPHICAL) WORLD AND THE PHENOMENAL WORLD

Except when introspecting or studying the behavior of groups, the psychological observer directs his attention to an individual organism or subject, *S*, and to the environment surrounding *S* (termed the "real" environment or, following Koffka, the "geographical" environment). Properties of *S* such as irritability, activity level, and docility can be determined by the observer. The observer can also determine such properties of *S*'s real environment as its spatial arrangement, the physical characteristics (elasticity, hardness, etc.) of its constituent parts, the nature and number of its floral and animal

---

*Received in the Editorial Office on January 29, 1959, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The author is grateful to D. K. Adams, J. J. Gibson, Fritz Heider, W. C. Langer, R. W. Leeper, and E. C. Tolman for their encouragement and for their helpful suggestions.

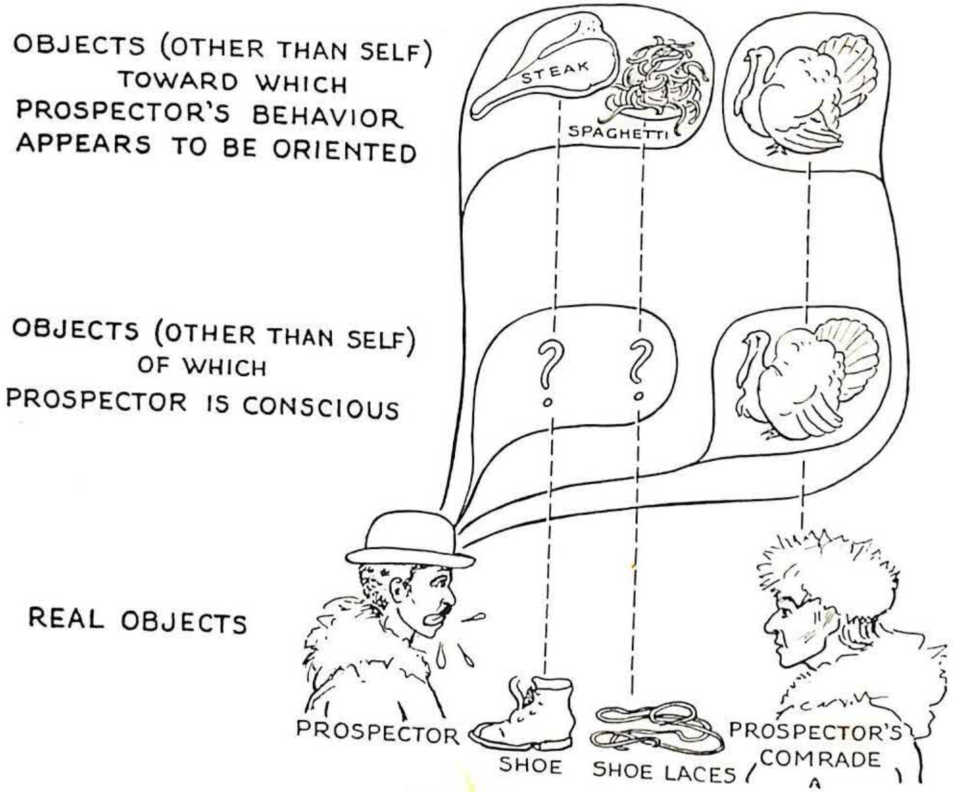

OBJECTS (OTHER THAN SELF) TOWARD WHICH PROSPECTOR'S BEHAVIOR APPEARS TO BE ORIENTED
STEAK
SPAGHETTI
OBJECTS (OTHER THAN SELF) OF WHICH PROSPECTOR IS CONSCIOUS
REAL OBJECTS
PROSPECTOR
SHOE
SHOE LACES
PROSPECTOR'S COMRADE
A

FIGURE 1
THE WORLD OF THE PROSPECTOR: REALITY AND FANTASY

inhabitants, and the constellation and dynamic inter-play of its human inhabitants. The existence of these properties of *S*'s real environment is usually largely independent of the fact of *S*'s existence.

The understanding and prediction of *S*'s behavior which is made possible by observing *S*'s interaction with the real environment can be enhanced by taking into account *S*'s immediate experience. The contents of *S*'s immediate experience, which constitute his "phenomenal" world, are made known to the observer through *S*'s description of the objects of his conscious awareness. Only when the observer serves as his own *S* can a phenomenal world become directly accessible to the observer. Although the features of the phenomenal world usually correspond to those of the real world, the phenomenal world follows laws of its own, to some extent. A given phenomenal world is distinguished from the corresponding features of the real world, in that it "belongs" to a given *S*, and is "carried along" by this *S*, to some extent, as he moves from one geographical locale to another.

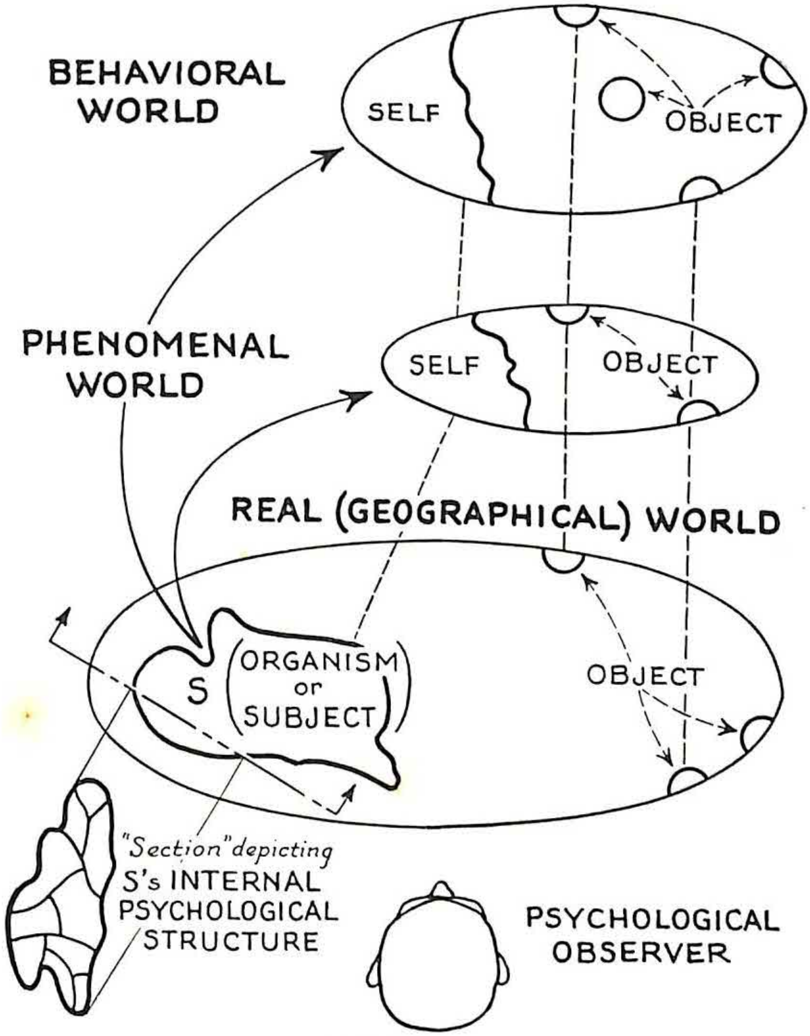


FIGURE 2
CONCEPTS ACCOUNTING FOR S'S BEHAVIOR

B. THE BEHAVIORAL WORLD

When *S* reports seeing an object which has no counterpart in the real world, he is said to be hallucinating. As the hunger of the "Gold Rush" prospector mounts, for example, he starts to chase his comrade with a hatchet. This behavior might be puzzling to the audience, except that the film-producer suddenly provides the audience with immediate access to the prospector's phenomenal world: the audience "sees," "through the prospector's eyes," that his comrade has turned into a gigantic turkey (see Figure 1), the only edible object he has seen for a while. The prospector must deceive himself in regard to his geographical environment before he can permit himself the indulgence of his growing cannibalistic wishes.

Such immediate access to the phenomenal world of the prospector is possible through magic, but not in everyday life. The prospector's self-report is the only avenue leading to his phenomenal world that is open to the psychological observer. In the absence of such self-report, the observer's desire to further his understanding of the prospector's hatchet-wielding behavior can be satisfied only by means of inferences about what is going on "inside the prospector's head."

Just such inferences had to be formulated to "make sense out of" the prospector's behavior in the "steak-eating" and "spaghetti-eating" sequences, since the film-producer provided no magic access to the phenomenal world of the prospector, in these instances. True, the audience understands the prospector's "steak-eating" and "spaghetti-eating" behavior *almost immediately*, since it realizes his desperate desire for food, and since Chaplin exaggerates the "steak-eating character" and the "spaghetti-eating character" of his movements so that the analogy will not be lost on anyone. Despite the rapidity with which the audience comprehends the nature of the objects to which the prospector directs his behavior, it must be kept in mind that "steak" and "spaghetti" are "constructs" developed by the audience to help account for the prospector's behavior; the audience has *no immediate* access to the prospector's phenomenal world in these sequences, nor is his self-report available. Whether or not steak and spaghetti are *truly* present in the prospector's phenomenal world is an unanswered question; it is conceivable that the prospector does *not see* steak and spaghetti, although he behaves *as if* he is reacting to these objects. Inferences regarding the "steak" and the "spaghetti" are therefore probabilistic.

Objects (such as the steak and the spaghetti referred to above) which are inferred from S's behavior are designated as "behavioral objects." The behavioral objects of a given S, taken together, constitute his "behavioral world" (see Figure 2). The behavioral world is defined as that world of objects around which S orders his behavior. A given behavioral world exists only in respect to a given S, and (as is also true for his phenomenal world) tends to be carried along by this S as he moves from one geographical locale to another.

S's behavioral world may correspond more or less closely to his phenomenal world, or to the real world. The phenomenal world is, in fact, a special case of the behavioral world, since self-report is one of the varieties of behavior on which inferences relating to the behavioral world are based.

C. "QUASI-CONSCIOUS" ASPECTS OF THE BEHAVIORAL WORLD

For ages past, unsophisticated observers have attributed to others behavioral worlds resembling the phenomenal worlds either of the *S*'s being observed, or of the observers themselves. Such "quasi-conscious" worlds have been ascribed to animals ("anthropomorphism") and to infants ("adultomorphism") by sophisticated as well as by unsophisticated observers. During the present century, scientific observers, reacting against the uncritical "phenomenologizing" of Romanes, have been unduly skeptical of the value of ordering *S*'s behavior in terms of the quasi-conscious aspects of his behavioral world—except, perhaps, in their extra-scientific efforts to understand and to regulate the wailing of their own infants or of their own pets.

Consider, for example, how the infant's behavior may be ordered in terms of the behavioral world.

The behavioral world of the *newborn* infant was formerly the subject of much speculation. (This speculation was actually carried on in terms of the phenomenal world, but the phenomenal world of the newborn is, strictly speaking, inaccessible to the observer, either directly or by means of self-report.) Since inferences relating to the structure of the behavioral world presuppose a greater capacity for vicarious, adaptive behavior than is possessed by the newborn, his behavioral world may be regarded as being highly undifferentiated, or possibly even non-existent.

The behavioral world of the two-week old is only slightly more developed. There is a minimum of coordination between the experiences mediated by the various senses, except in the region of the mouth (the principal organ for manipulation, grasping, and incorporation), where a primitive distinction between "self" and "non-self" has been formed. The behavioral world at the age of two weeks is represented as a series of discrete, undifferentiated "sensory sectors" (see Figure 3), except for the "original" space of the mouth.

Considerable change has taken place in the behavioral world by the age of six weeks. The developing capacities for ocular pursuit and for voluntary movement of arms and hands permit a coordination of visual and tactual-proprioceptive experiences of the mouth and hand, with a resulting extension outward from the "original" space of the mouth. As differentiation proceeds within the various sectors, and as these sectors become better coordinated, the organization of the behavioral world in terms of sensory sectors comes

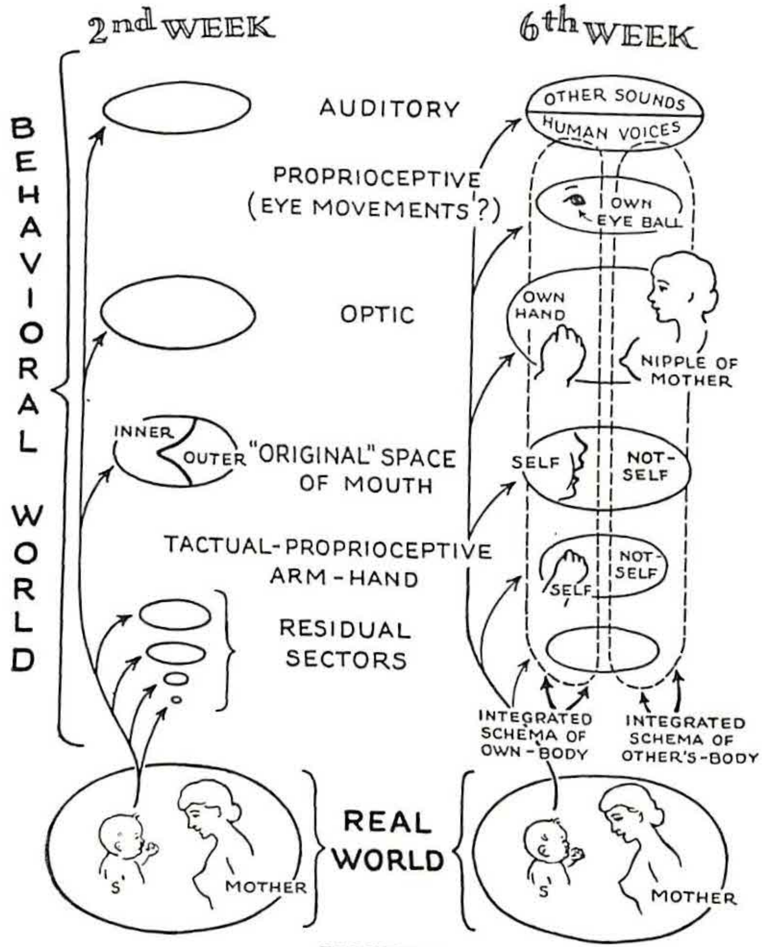


FIGURE 3
DEVELOPMENT OF THE QUASI-CONSCIOUS ASPECTS OF THE BEHAVIORAL WORLD

to be supplemented, and partly supplanted, by its organization in terms of "self" and "non-self" objects, which may be represented in several sectors.

Nuclei for the "schema" of the six-week old's own body are established and developed by the tactual-proprioceptive experiences of his mouth and his hands, by the sight of his hands, by sensations from his intestinal tract, by the sound of his voice, and possibly even by the sensations produced by his tears and by the movements of his eyeballs. The "schema" of his body is confirmed, clarified, and extended by his growing capacity to observe the bodily features and movements of others.

"Non-self" objects whose real counterparts provide more variable stimulation than that provided by *S*'s own body also begin to stabilize in the behavioral world of the six-week old. Mother is regarded as still the same person, even though her visual aspect may vary from moment to moment. Even when she is temporarily absent, "mother" continues to exist in the six-week old's memory, and he anticipates her return.

D. "UNCONSCIOUS" ASPECTS OF THE BEHAVIORAL WORLD

The observer constructs behavioral worlds as quasi-conscious surrogates for phenomenal worlds to which he lacks access. This by no means exhausts the functions served by the behavioral world.

There are occasions when *S*'s behavior is best understood in terms of his reactions to objects which are not only beyond the direct awareness of the observer, but are also beyond the direct awareness of *S* himself. At such times, the observer may postulate one or more behavioral objects, when it seems that *S* is acting *as if* these objects are present, even though these behavioral objects may have no immediate phenomenal or real counterparts. Just such "as-if" objects were postulated, it will be recalled, to help "make sense out of" the prospector's behavior.

An ample appreciation of the potentialities of the behavioral world as a conceptual tool for the observer is evident in the psychology of Freud. The behavioral world, although not formulated explicitly by Freud, plays a central rôle in the psychoanalytic account of behavior. The phenomenal world is, by contrast, reduced to somewhat of an epiphenomenon, regulated by repressive and other internal "mechanisms" whose state is, in turn, closely related to that of the behavioral world.

The fruitful rôle assigned by Freud to the behavioral world is perhaps nowhere better indicated than in Freud's account of *S*'s reaction to the loss of a love object, in mourning and in depression. Confronted by the death of a person he loves, *S* may at first find it difficult to "get himself to believe" what has happened, and he may occasionally hallucinate the loved person. While death occurs almost instantaneously in the real world, the loved person may persist as an object in *S*'s behavioral and phenomenal worlds, undergoing rather gradual changes during the course of mourning. The representation of the deceased love object in the behavioral world of the mourner undergoes some distortion, as by romanticization. More importantly, however, the representation of the deceased usually comes to be reduced, occupying a less prominent rôle in *S*'s behavioral world, with the result that the discrepancy between the behavioral world and the real world, brought about by the death of the loved person, is lessened.

Identification is, in Freud's view, a principal means for reducing the discrepancy between the behavioral world of the mourner and the real world. *S*, in identifying with the deceased, takes on some of his internal psychological structural features. A corresponding change occurs in *S*'s behavioral world, where the representation of *S*'s own self is enhanced at the expense of the representation of the deceased. The change in the behavioral world, which can be regarded as an "incorporation" of some aspects of the deceased into the self, is likened, in Figure 4, to the ingestion of a foreign body by an

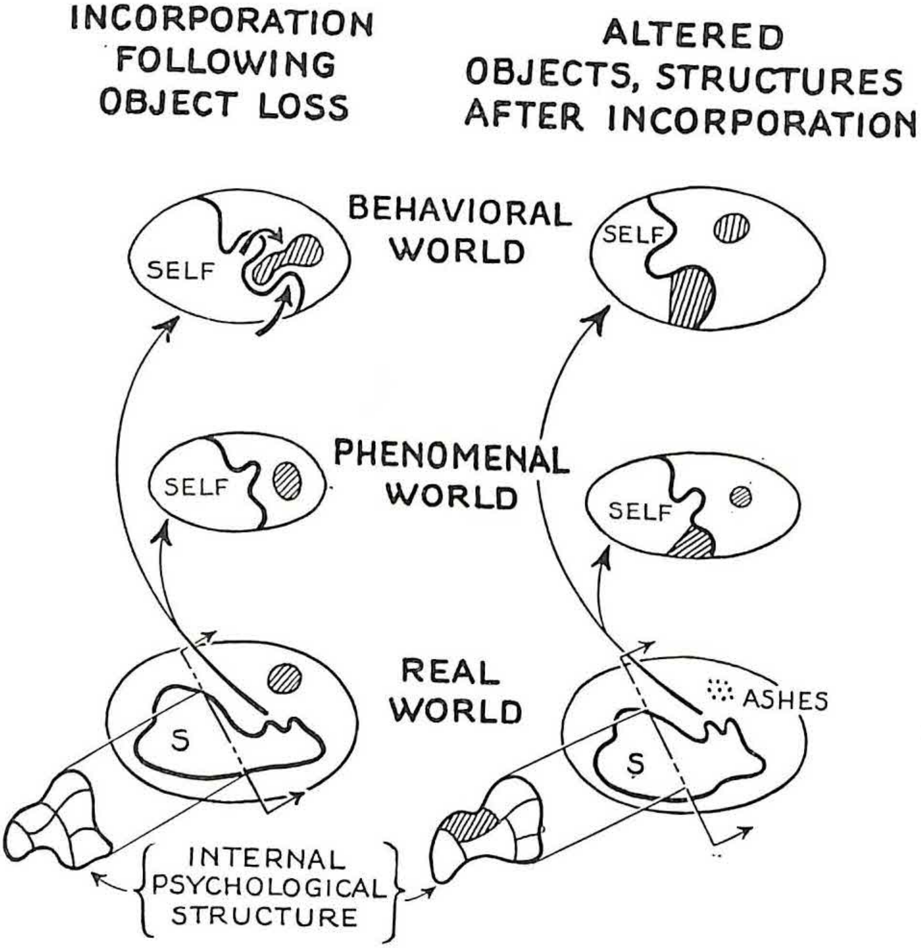

INCORPORATION FOLLOWING OBJECT LOSS
ALTERED OBJECTS, STRUCTURES AFTER INCORPORATION
BEHAVIORAL WORLD
SELF
SELF
PHENOMENAL WORLD
SELF
SELF
REAL WORLD
S
S
ASHES
INTERNAL PSYCHOLOGICAL STRUCTURE

FIGURE 4

INCORPORATION AND IDENTIFICATION IN MOURNING

Shading denotes: 1) physical presence of deceased person; 2) persisting representatives of deceased person; and 3) aspects of self, or of internal psychological structure, altered in identification.

amoeba. The ingested material is broken down and then re-formed into new structure which, while it is assimilated into a new entity, retains traces of its former identity. In this way, *S* simultaneously preserves something of value from the past, while coming to terms with current realities—transforming himself, as it were, into a living memorial to the deceased.

An important rôle is assigned by psychoanalytic observers to incorporation of lost love objects, in psychopathology and in normal development, as well as in mourning. Incorporation of a lost love object is said to play a rôle in depression similar to that in mourning, except that it is prompted by "psychological loss" rather than by loss through death. Marked depression may follow prolonged absence of the mother at a critical phase of childhood, for example. Melanie Klein, elaborating on Freud's suggestion that the "superego" is formed during the normal course of development as a "precipitate" of the child's earlier relations with his parents, has proposed that the superego is laid down by successive incorporations and projections, beginning in the earliest weeks of life. Disturbances in these early transformations are held responsible by Klein for various structural weaknesses predisposing to psychopathology in later life. The very differentiation between "self" and "non-self" objects, referred to earlier, is dependent, in Klein's view, on the earliest of these incorporations and projections.

E. INTERNAL PSYCHOLOGICAL STRUCTURE AND THE BEHAVIORAL WORLD

The present discussion began with a consideration of *S* vis-a-vis his geographical environment. In the hope of adding to the observer's understanding of *S*'s behavior, the behavioral world—a hypothetical structure conceived as being external to the boundaries of *S*—was introduced. It would have been possible, alternatively, to account for *S*'s object-oriented behavior in terms of a hypothetical structure conceived as being localized *within* the boundaries of *S*. It is, indeed, more usual for contemporary psychologists to think in terms of internal psychological structure than in terms of the behavioral world. It will be recalled, however, that Freud used both concepts simultaneously, suggesting that changes in the behavioral world might be accompanied by changes in internal psychological structure. This raises the question: How are internal psychological structure and the behavioral world related to each other?

Several kinds of relation are possible between internal psychological structure and the behavioral world.

Some observers have felt it wisest to place the burden of psychological explanation on a single kind of construct alone.

McDougall relied exclusively on structures attributed to *S*'s "interior," in attempting to account for behavior. Impressed by *S*'s activity, and by the ordering of this activity around objects in the real world, McDougall conceived of behavior as being initiated by internal sources of energy ("propensities") and as being directed by internal structures ("sentiments"). The sentiments are elaborated slowly, and eventually match the geographical objects to which they correspond, as a glove matches a hand. Given such a close match between internal structure and external (geographical) object, McDougall felt no need to think in terms of objects other than those belonging to the real world.

In contrast to the active, vicarious behavior emphasized by McDougall is the "imprinted" behavior studied by Lorenz and Tinbergen, in which *S* follows slavishly in the footsteps of objects belonging to the geographical environment. At first glance, it might seem that the imprinted *S* is "pulled along" from without, like a magnetized needle in the field of a strong magnet. Despite *S*'s "passivity," Lorenz and Tinbergen would probably hold that the imprinted behavior is maintained largely by energies and structures internal to *S* (these structures would, of course, be simpler and more rigid than those postulated by McDougall). Lorenz and Tinbergen would probably feel as little need as McDougall to invoke the idea of a behavioral world.

While sociologists tend to account for behavior primarily in terms of *S*'s *real* environment, it is difficult to cite instances of observers who place the burden of explanation for *S*'s behavior solely on his behavioral world. Lewin is usually regarded as the principal advocate of explanation in terms of the behavioral world, but, as will be seen, he relied on other concepts as well.

The possibility of accounting for behavior exclusively in internal terms or exclusively in external terms suggests that internal psychological structure and the behavioral world may be interchangeable concepts. If this is true, one concept or the other is superfluous, and the choice between them is arbitrary. If, on the other hand, each concept is best suited for handling somewhat different sets of observations, the two concepts are not simply duplicates of each other, but may be complementary, for example; or overlapping, but not interdependent. Various observers have made use of both kinds of concepts, assigning somewhat different functions to each.

Freud, for example, sometimes dwelled on internal psychological structure, at other times on the behavioral world; at still other times, he employed both concepts simultaneously. Since Freud's energies were devoted more to exploring new problems than to examining systematically the logical relations between the constructs he employed, it is not surprising that the relations between internal psychological structure and the behavioral world which he depicted are extremely rich and suggestive, but not too clearly defined.

Murray, attempting to integrate the views of Freud, McDougall, and others, employed an internal structural element, "need," as his core concept. Murray related needs to specific external objects, as did McDougall, by introducing a complex internal structure, the "need-integrate" or "sentiment." At the same time, the "press"—an external construct related to the behavioral world—was introduced to account for various psychoanalytic observations. Murray's attempt at systematization must be regarded as programmatic, however, since two distinct approaches—one independent of or neutral with respect to the behavioral world, the other associated with the behavioral world—are incorporated into a single system of thought, without being assimilated or otherwise related to each other.

Perhaps the most interesting suggestions for relating internal and external constructs to each other and to behavior are those proposed by Lewin. Behavior (or "locomotion") is, according to Lewin, a resultant of the psychological "forces" acting on S. A given psychological force, f, is co-determined by the "demand-character" or "valence," Va, of one or more objects in S's "life space" (the "life space" is a concept similar to the "behavioral world"), and by the distance, e, between S and these objects. Demand-character is, in turn, co-determined by the "tension," t, of internal sub-systems, and by non-psychological, "alien" factors, G, (i.e., features of the geographical environment). These suggestions, expressed by Lewin as follows

$$\mathrm{f} = \mathrm{F}\left(\frac{\mathrm{Va}}{\mathrm{e}}\right) = \mathrm{F}\left(\frac{\mathrm{t,G}}{\mathrm{e}}\right),$$

together with Lewin's proposals for anchoring the various concepts firmly in behavior, constitute the principal attempt, to date, to formalize the relations between internal psychological structure, the behavioral world, and behavior.

It appears, on the basis of this brief survey, that internal psychological structure and the behavioral world differ somewhat, both in regard to the events of the real world which they represent, and in regard to the manner in which they represent these events. A categorical statement of the precise differences between these two concepts would be premature, since terms referring to either concept are too frequently obscure or else vary in meaning from observer to observer. Of the alternative formulations which are possible, however, some such multi-stage scheme as that advanced by Lewin—with internal psychological structure being defined in terms of behavior and in terms of other real events, with the behavioral world being defined in terms of internal psychological structure and in terms of still other real events, and with behavior being related to the behavioral world and to yet other real events—some such scheme appears to the writer to hold promise.

F. CONCLUDING REMARKS

The behavioral world is an entity invoked by the observer to help account for *S*'s behavior. While accounts of behavior presented in terms of the quasi-conscious aspects of the behavioral world, or in terms of its unconscious aspects, have often excited the imagination, these accounts have heretofore been largely intuitive. For the behavioral world to become a useful instrument for the scientific study of behavior, it is necessary that we obtain a fuller understanding of the functions which can be served by this concept.

Careful explorations, such as that undertaken by Lewin, are called for, to determine the various ways in which the behavioral world can be formulated, the various ways in which it can be related to internal psychological structure, and the various operations by which these two concepts can be coordinated, without ambiguity, to events in the real world. It is hoped that, through such explorations, ways can be found to characterize the behavioral world which are independent of the whims of the individual observer, and free from circular reasoning. It is further hoped that the behavioral world can be so formulated as to foster the discovery of a body of empirical laws capable of serving as a basis for the prediction of behavior.

Department of Neurology & Psychiatry
Northwestern University Medical School
303 East Chicago Avenue
Chicago 11, Illinois

Published as a separate and in *The Journal of Psychology*, 1959, **47**, 289-295.

COMPARISONS OF PHENOMENAL DISTANCE IN PHOTOGRAPHS OF VARIOUS SIZES*

Department of Psychology, Michigan State University

S. HOWARD BARTLEY AND HARLEY J. ADAIR

A. PROBLEM

As a part of an extended study of the phenomenal spatial features of pictorial material, in which two studies (1, 3) have already been completed, the present paper reports the relation between sizes of photographic prints and their viewing distances in order that material in them appear to be equally far away. This information has been obtained not by direct questioning about apparent (phenomenal) distance but rather by a psychophysical procedure in which the same material in several sizes of prints is matched for some property such as apparent distance. The task of the observer is to adjust the metric distance of the larger of the two prints until the required match between the material in it and in the smaller reference print is accomplished. The metric or viewing position of the smaller of the two prints is fixed for a given set of comparisons, but varied from set to set. From this procedure, the relation between the metric distance, and metric size of the larger of two prints and the metric distance of the smaller of the two prints is obtained.

B. MATERIALS AND PROCEDURE

The target material for the study consisted in a small reference print 1″ × 1″ (Figure 1) and, three large comparison prints of the same scene, 3″ × 3″, 4″ × 4″, and 6″ × 6″ respectively. The viewing distance of the small fixed print was, from time to time, at 16″, 24″, and 32″.

To present the prints a calibrated track about 22′ in length was used. This track supported a carriage for the large print. The turning of a crank moved the carriage to the desired point along the track. The small print was mounted on a holder alongside the track, and at the appropriate height. The carriage provided for seeing both the large and the small prints side by side at eye level. Moving the large print toward or away from the viewer did not raise or lower it in the visual field. The two prints were seen against a homogeneous background of low reflectance ("black") in a diffusely illuminated room.

---

*Received in the Editorial Office on January 30, 1959, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

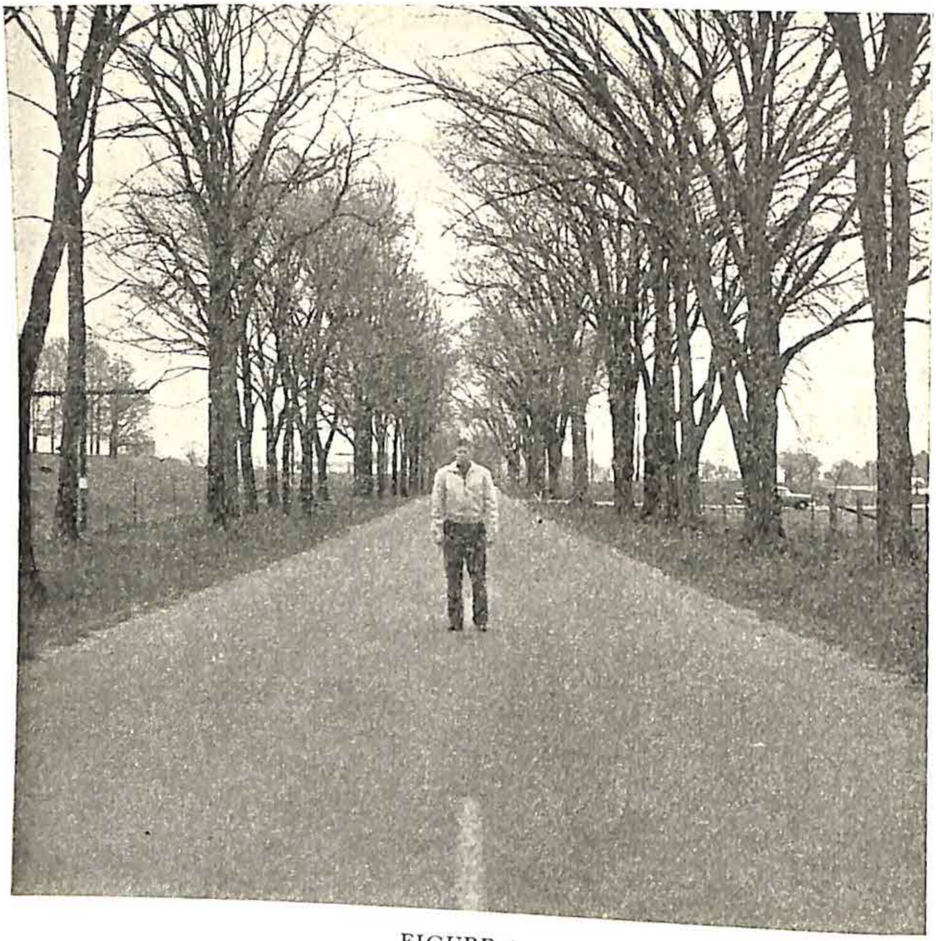


FIGURE 1

The usual kinds of procedural precautions were used so as to minimize the influence of extraneous factors. Fifteen observers were used all of whom were unaware of the purpose of the study and rather untrained in observing. This sort of a group was chosen not because of a supposed inappropriateness of trained observers. It was supposed, among other things, that if an untrained group would provide the study with orderly results, it might be said that such results would more nearly represent what could be expected of the population at large. Trained observers might be expected to yield data with less spread (variability) than untrained.

C. Results

The results are pictured in Figure 2(*a*) in which the mean metric distance of the large print from the eye is plotted against its metric size (3″, 4″, or 6″). Curve *A* is the end-result when the small references print is 32″ from the eye; Curve *B*, when the print was 24″ from the eye; and Curve *C*, when it was 16″ from the eye.

Figure 2(*b*) using the same data as for Figure 2(*a*), shows the viewing distances of the larger targets as dependent upon the viewing distances of the small one. Each curve is for a different size large target. In Figure 2 (*a*), it can be seen that the over-all mean distance of the large print (print varied in size, and distance) becomes greater as the metric distance of the fixed print increases. Also, as the large print size increases, its mean distance increases accordingly. And further, the slopes of the three curves differ; the

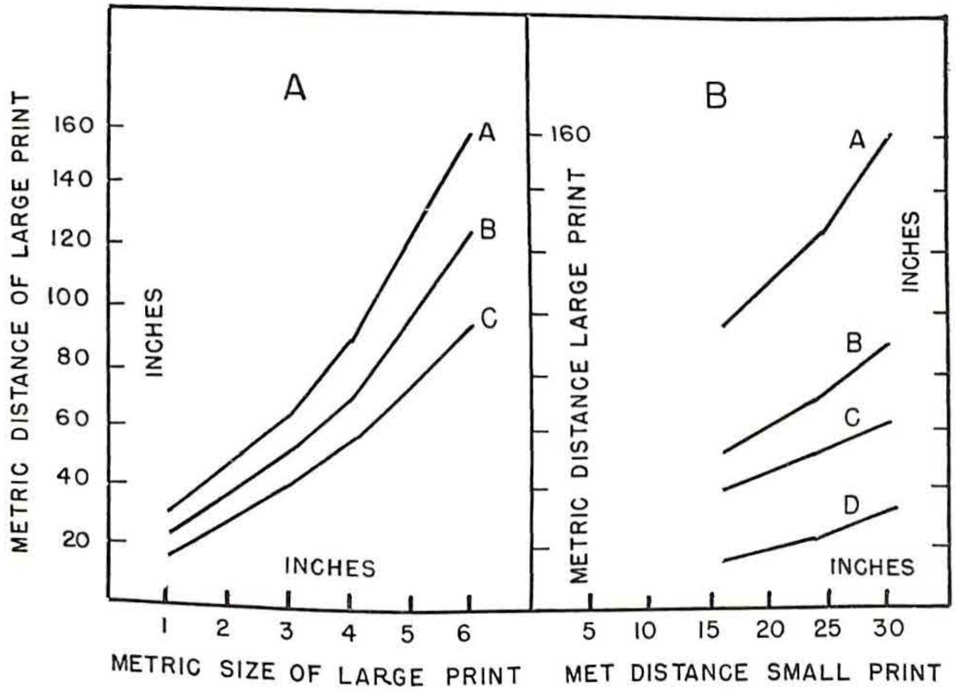


FIGURE 2

(*a*) The relation between metric (viewing) distance of large print (varied in size) and the metric size of these prints. Curves *A*, *B*, and *C* are for viewing distances of 32 inches, 24, and 16 inches, respectively for the small print of fixed position. The large print was varied in distance to match the small fixed print in apparent distance.

(*b*) The relation of metric distance of large print to metric distance of small print. Curves *A*, *B*, *C*, and *D* are for large print sizes of 6, 4, 3, and 1 inches, respectively. The points on the graphs for the large print when it is a 1 inch square, are presumptive, i.e., supposed to be the same as the viewing distance of the small print.

16″ distance curve has the smallest slope, the 32″ the largest, with slope of the 24″ distance curve falling between them. This is indicative of an interaction between distance of the fixed small print and the distance of the large print and also its size. That is, increasing the large target size may have resulted in mean distance of large print increments which were greater for fixed print distances of increasing magnitude.

The results of an analysis of variance performed on the scores of the 15 *S*s are reported in Table 1. Each *S* was given 54 trials, six for each of the nine distance and size combinations. There were, then, a total of 810 scores obtained.

A Bartlett's (2) test for homogeneity of variances led to the rejection of the hypothesis of homogeneity of variances ($F = 9.01$, p. .01), but since the mean differences in scores obtained for the several groups were observed to be great, and therefore, the expected *F*-values were large, an analysis of variance was performed with the expectation of following Lindquist's (4) suggestion. It is to the effect that when marked heterogeneity of variance is found, "it is desirable to allow for the discrepancy by setting a slightly higher 'apparent' level of significance for the test than one would otherwise employ (the 'apparent' level being indicated by the *F-table*)" (p. 83). This procedure was followed, and since the *F*'s reported in Table 1 are all considerably greater than the tabled *F*-values (for appropriate *df*) at $p < .01$, it may be said that the *F*'s obtained are significant at $p < .01$.

The means for print distance are significantly different at $p < .01$, with $F = 58.71$, at *df* 2,28, and likewise for mean differences for print size ($F = 56.86$, with *df* 2,28). The interaction between $D \times S$, suggested previously, is also seen to be significant at $p < .01$, with $F = 10.53$, at *df* 4,56.

TABLE 1

ANALYSIS OF VARIANCE OF MEAN DISTANCE OF VARIABLE (LARGE) PRINT IN INCHES FOR CONSTANT (SMALL) PRINT DISTANCE AND VARIABLE (LARGE) PRINT SIZE

Source	Sum of squares	*df*	Mean sq.	*F*
Distance (*d*)	226,071.65	2	113,035.82	58.71**
Size (*S*)	839,602.69	2	419,801.34	56.86**
$D \times S$	42,806.12	4	10,701.53	10.53**
Observer (*O*)	371,239.10	14	26,517.08	516.20**
$D \times O$	53,912.76	28	1,925.46	37.48**
$S \times O$	206,730.83	28	7,383.24	143.73**
$D \times S \times O$	56,906.58	56	1,016.19	19.78**
Error	34,676.67	675	51.37	
Total	1,831,946.40	809		

**Significant at $p < .01$.

Increasing large target size results in relatively different increases in mean distance of large target for increasing small fixed print distances.

The between-observer means are also significantly different at a high level of confidence ($F = 516.20$, $p < .01$, for df 14,675). Observer's judgments of equidistance of fixed and variable targets differed greatly. The range of over-all mean distance judgments for observers (mean of all judgments made by each observer) was, in fact, 59.15 to 116.44 inches. For example, Observer 2, on the average, judged a large variable print distance of nearly *twice* the mean distance judgment of Observer 14 as being equidistant with the fixed small print.

The significant interaction of $D \times O$ ($F = 37.48$, $p < .01$ for df 28,675) may briefly be explicated as follows: For different observers, increasing the fixed print distance from 16 to 24 to 32 inches resulted in judgments of differing relative sizes. For example, changing fixed print distance from 16 to 24 inches may, in one observer, result in a marked change in distance judgment (placement of variable print), while in another observer the corresponding change may be relatively small.

A similar interpretation may be made for $S \times O$ interaction ($F = 143.73$, $p < .01$ with df 28,675). Changing size of large print resulted in changes in judgments of relatively different amounts for different observers.

Finally, the significant $D \times S \times O$ interaction ($F = 19.78$, $p < .01$ with df 56,675) indicates that interaction effects of the highest order occurred.

The results of the analysis of variance, to reiterate briefly, indicate that significant differences in means for simple effects, D, S, and O were obtained and that, in addition, higher order effects, $D \times S$, $D \times O$, $S \times O$, and $D \times S \times O$ were likewise statistically significant. That is, not only were the simple effects significant, but interactions of all combinations of simple effects were also statistically significant.

D. Discussion

The results are a demonstration that the phenomenal distance of elements of a scene vary in keeping with the size of the print viewed and in keeping with its metric distance from the eye. This needs to be viewed in relation to what might be called distance constancy. Whereas size, shape, color, and lightness constancies have been quite extensively studied, distance constancy has received very little attention. Accordingly we do not know so much about it, or what might well be meant by the term, distance constancy, itself.

One form of distance constancy was studied by Smith (5) in which a certain marked-off extent at some distance from the observer was compared to a reference "extent" or distance nearby. Constancy under these conditions was reported. But although this situation involved, as a crucial aspect, what was called distance constancy, the variables studied were actually quite unlike those studied in the present investigation.

It is common in present-day thinking to believe that perception of objects follows one or the other of two laws, or else a compromise between them. The first is the law of the *visual angle.* In which case, two targets of different dimensions would look equal if the larger of the two were placed far enough beyond the smaller for it to subtend the same visual angles as the smaller one. The second law is the law of *constancy,* in which case two targets of differing dimensions would look equal in size regardless of their distance from the eye, and thus regardless of visual angles subtended. According to the usual way of looking at the matter, if the results turn out that the required distance for the larger target is somewhere between the point required by the law of the visual angle and the same distance as the smaller of the two targets, a compromise between the two laws is demonstrated. According to our way of looking at the matter, how can this be a compromise? How can a numerical compromise between a fixed value, and an infinite number of values all of which are equally good (constancy) occur?

In the present case the points on the curves of distance in (*a*) of Figure 2, represent, for the most part, a certain narrow range of percentages (mostly 76-88 per cent) of what would be expected according to the law of visual angle, but *not* the visual angle itself. Whatever the influences are that oppose the visual angle being strictly followed (for apparent distance), the law of constancy (sheer numerical indifference) would not seem to be a salient one.

E. SUMMARY

The phenomenal distance of an object in three large prints was matched with that of the same object in a standard small print of the same scene. The standard was placed in each of three fixed distances. It was found that the larger the print the further away it had to be for the match to result. Also the further away the small reference print was placed, the further away the large prints had to be placed. The slopes of the curves representing these relations differed in the three cases.

REFERENCES

1. ADAIR, H. J., & BARTLEY, S. H. Nearness as a function of lateral orientation in pictures. *Percept. & Mot. Skills,* 1958, 8, 135-141.

2. BARTLETT, M. S. Tests for homogeneity of variance. *In* Dixon, W. J., & Massey, F. J. (*Eds.*), *Introduction to Statistical Analysis.* New York, McGraw-Hill, 1957.

3. BARTLEY, S. H., & THOMPSON, R. A further study of horizontal asymmetry in the perception of pictures. (In manuscript.)

4. LINDQUIST, E. F. Design and Analysis of Experiments in Psychology and Education. Boston: Houghton-Mifflin, 1953.

5. SMITH, O. W. Distance constancy. *J. Exper. Psychol.,* 1958, **55**, 388-389.

Department of Psychology
Michigan State University
East Lansing, Michigan

Published as a separate and in *The Journal of Psychology*, 1959, **47**, 297-304.

MMPI SCREENING OF ENTERING MEDICAL STUDENTS*

Department of Psychiatry of Cornell University Medical College and Department of Psychology of Hunter College

CHARLES A. KNEHR AND RICHARD N. KOHL

A. PROBLEM

Psychiatric counselling and treatment have been available to students of the Cornell University Medical College for a period of 22 years (8). However, the traditional admissions procedures have not been specifically designed to evaluate or predict the presence or probable future appearance of emotional and adjustment problems in new students, either during or following medical education.

Starting with the class of students entering in 1951, a study was initiated for the purpose of relating admissions data, supplemented by a quantitative personality inventory, to the observed adjustment of students during the subsequent years of medical training.

Primary interest specifically centered in the problem of whether those students who later availed themselves of psychiatric help could have been detected at the time of admission. A small minority of students drop out each year for psychiatric and other reasons. Interest here was in the motivational or other personality characteristics which might differentiate this group of drop-outs. One could not attempt to study the factors which make for maladjustment in medical students without being concerned with the search for clues as to the characteristics of the successful medical student, and the physician who evidences superior performance in some area of medical practice or research. Since other admissions data were simultaneously available, an additional aim was to check the predictive value of college grades and aptitude test scores.

The present study is concerned with several facets of the extremely complex problem of efficient medical student selection. A most complete review of the many considerations is provided by the published report of the Fourth Teaching Institute of the Association of American Medical Colleges (4). Therefore no review of current thinking and pertinent literature is undertaken here.

---

*Received in the Editorial Office on January 31, 1959, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

B. PROCEDURE

The admissions records include results of a personal interview, personal data, and a moderate amount of family history as recorded on the application form, letters of recommendation, a transcript of academic performance in college and, for 96 per cent of entering students, the grades achieved on the four parts of the Medical College Admissions Test (*MCAT*).

Within one month of beginning the first year in medical school, five successive classes of entering students were given the group form of the Minnesota Multiphasic Personality Inventory (*MMPI*). The tests were scored, coded, and profiled, but the results were not made available to any member of the faculty.

The first three of the five classes which were tested initially have completed medical school training ($n = 249$) with complete information available as to yearly and cumulative class standings, the data of those who failed to complete their medical education, with the stated reasons for purposes of record, and psychiatrists' notes on the students in these three classes who received consultation or treatment.

Possible relationships between the quantifiable data were examined by means of plotted correlation diagrams and any relationships, even remotely suspected, were then calculated as Pearson r's.

In the case of *MMPI* profiles sorted with respect to abnormality, absolute statistical elevation of one or more scales was used as a criterion on the first approximation. However, in the results reported below, the profiles indicated as abnormal were coded and sorted in accordance with profile characteristics that clinical experience indicates is important from the standpoint of overall adjustment. The judgments as to normality or abnormality in the profiles were made by the senior author without the knowledge of the students' academic performance or adjustment level.

C. RESULTS

The number of *MMPI* profiles suggesting significant psychopathology in entering medical students was 63 out of 249 students (25.3 per cent) from the three classes that have now been graduated. These three classes also yielded 37 students who were known to have been treated for circumscribed psychiatric illnesses while in medical school. However, only 18 of these treated students showed abnormal profiles at the time of beginning their medical training. In effect, over two thirds of the students with test indications for psychopathologic functioning were not known to have sought psychiatric help, nor did they attract the attention of faculty members because of psychopathologic behavior. Approximately half of those students treated during the four years of medical school did not reveal personality aberrations at the time of testing.

Of the same 63 students with abnormal profiles, those who did and did not request psychiatric help could not be separated on the basis of the frequency or height of profile elevations, or on the basis of diagnostic indications. The special control scale (*Cn*) developed by Cuadra (2) for the *MMPI* did not distinguish relative adjustment among the students with abnormal profiles, the raw score means for treated and untreated students being 27.9 and 28.0 respectively. While biased somewhat by five students in this group who dropped out and whose low standings therefore did not appear in the mean figure, the mean cumulative class standing of students with abnormal profiles was not significantly different from that of students who showed no clear test indications for psychopathologic features.

Correlations were sought between cumulative class standings as the available criterion for success in medical school training and the separate scales of the *MMPI* including the validating scales, the status scale (6), and the control scale. No significant correlations were observed in a class taken as a whole. Total profile elevation yielded similar negative results.

Some trends were noted when a comparison was made between the 15 highest, average, and lowest passing students in terms of final cumulative class standings, and the drop-outs. It should be noted that these relationships are without statistical significance in the student group as a whole. Table 1 shows an increasing incidence of elevated scores on the Lie scale as students of the subsample rank increasingly lower. This indicates a tendency to heightened levels of unrealistic self-appraisal in the less competent students of these sub-samples.

The mean *MMPI* profile codes for these same four groups of students are shown in Table 2. The trend is interesting though not statistically significant. The least successful students tend to show a slightly greater frequency of moderate elevation of scales related to psychoneurotic features whereas the successful students tend more toward non-empathic, hypomanic characteristics

TABLE 1

Group	Number	L scores of 5 and above	Per cent L scores above 5	Mean L scores
Highest	15	3	20%	3.1
Middle	15	6	40%	3.5
Lowest	15	10	67%	5.1
Drop-outs	9	6	67%	5.2

TABLE 2

Group	Welsh code (10)
Highest	5-49378621/
Middle	5-43789126/
Lowest	53-7894621/
Drop-outs	3-54187926/

as indicated by their profiles. Looking at the profiles individually, Scale 3 was one of the two highest scales in seven out of nine drop-outs while this occurred in only six of 15 highest students. A high point of Scale 9 occurred only once in drop-outs and in six of the top students.

Since such personality characteristics as were observed could hardly be said to be crucial in the adjustment or academic performance in three classes of medical students, undergraduate grade-point indices and the *MCAT* scores were scanned for possible relationships with success in medical school. For the student group as a whole, there were significant but low correlations of .40 and .41 between the *MCAT* science scores, the mean *MCAT* scores, and final cumulative standings of students. The *MCAT* verbal, quantitative, and social scores showed significant but lower correlations of .31, .24, and .36 respectively with final standings.

Since the *MCAT* science subtest tended to predict grades better than the other *MCAT* subtests, the distribution of the *MCAT* science scores for the four subgroups of highest, average, and lowest passing students, and drop-outs is shown in the form of a contingency table (Table 3). The ranking of these students is clearly related to achievement levels on *MCAT* science. In relating the reasons for withdrawal from medical training, the most capable students appear to have failed for reasons of a psychiatric nature. The remaining drop-outs were recorded as having withdrawn for various non-psychiatric reasons such as lack of interest in medicine. The tabulation suggests that an inadequate achievement base, an unsuitable ability configuration, or insufficient

TABLE 3

		MCAT science score					
Group	N	400 to 449	450 to 499	500 to 549	550 to 599	600 to 649	650 and up
Highest	15		1	1	3	4	6
Average	15	1	1	2	6	2	3
Lowest	12*	1		7	4		
Drop-outs	8*	2	2	1	1**	1**	1**

*The *N* is reduced by students not having *MCAT* scores.
**Withdrew for psychiatric reasons.

interest is at least as important in this group of drop-outs as is the problem of adjustment.

A related trend is noted in these four sub-groups in the grade-point ratios in college science and humanities courses (Table 4). There were large variations in the numbers of science courses that medical students completed as undergraduates, and in the ratio of these to the number of humanities courses.

TABLE 4

Group	Mean undergraduate grade-point ratios		Δ Science grades
	Science	Humanities	
Highest	3.55	3.26	.29
Average	3.26	2.99	.27
Lowest	3.13	2.99	.14
Drop-outs	3.19	3.31	—.12

There is a trend to lower undergraduate science grades in lower ranking students, but the differences are not significant. Of some interest, but again not statistically significant, is the amount by which the mean science grade-point ratio exceeds that for the humanities in the four sub-groups. The higher the student's undergraduate science grades relative to grades in the humanities, the more he tends to stand high in his class. This finding is apparently not altered by the difference in approach which might be required by basic science and clinical subject matter taught in the first and fourth years of medical school respectively. Top students showed an average standing of 6.9 in the first year and 4.8 in the fourth, while the mean standing for the lowest students in these years was 75.9 and 77.0. In the case of the drop-outs, there is a further slight trend to excel in the humanities as compared to science subjects. In the total student sample, however, these trends are no longer evident.

The incidence of physician fathers among medical students is sufficient (14.8 per cent in this sample) to invite inquiry as to the relative effectiveness of these students in medical school training. The final cumulative class standing of these students was not significantly different from the mean standing of students whose fathers' occupations vary from restaurateur and engineer to college professor.

D. DISCUSSION

Upon initiation of this study, it was assumed that accepted applicants to medical school would have been so selectively screened during 15 or 16 years of schooling prior to admission that, for all practical purposes, the level of general mental ability and prior educational achievement would not be crucial factors in successful medical training. Motivational and personality factors would seem, a priori, to play the more significant rôle.

Psychological appraisal of non-intellectual characteristics has seen limited use by the various medical schools. Our choice of the *MMPI* was for the purpose of screening incoming students for the presence of psychopathologic features, and not to study personality variables in general in relation to successful medical training.

The *MMPI* suggested clinical levels of psychopathologic functioning in approximately a quarter of beginning medical students. The lack of correlation of the *MMPI* clinical scales with class standing, and the finding that less than a third of the students with abnormalities in the *MMPI* profiles sought or appeared to be in need of counselling or treatment, suggests that psychopathology is only one of a matrix of factors in successful completion of medical training. If we assume that the test was a valid indicator of deviant personality trends, then it is clear that the presence of psychopathology as indicated by test in individual students cannot be used to predict future adjustment to a sufficiently practical degree to recommend incorporation of the test into regular admissions procedures.

In reference to predicting which students would need psychiatric assistance, half of the treated students provided no test indication at the time of admission of observed later adjustment difficulties. The other half of the treated sample were among those with abnormal profiles, but they could not have been detected at the time of admission from the information available.

Not until a smaller sample of the best, average, and lowest ranking students was drawn from the total sample of three medical school classes did a few trends become evident in the profiles. These are in partial agreement with previously reported studies using the *MMPI* (5, 9). Among these were the increased incidence of high Lie scores (egocentrism) in drop-outs and lowest ranking students, and a higher incidence of elevated *Hy* (conversion type symptoms) scores in the drop-outs, although in only one student did the scale reach a standard score of 70. Again it should be noted that for the entire student sample, these relationships did not hold.

In studying other admissions data for their possible predictive value in these three student classes, several significant relationships were shown. Low but significant correlations were found between the *MCAT* scores, the mean *MCAT* score, and cumulative class standings. In relation to the other data of this study, these findings lend support to the conclusion reached by Funkenstein (3) that the best present predictors of successful achievement in medical education are those of the Medical College Admissions Test.

The spread of ability and achievement levels in our students as indicated by the *MCAT* and by undergraduate college grades was larger than expected. This observation seems in agreement with the statement of Wolfle (11) to the effect that, because students frequently choose other fields in preference to medicine, medical selection committees never get the first choice of graduate level students. Further, from among those who do apply to medical school, an admissions committee selects a top group and alternates, many of whom choose other schools. Each year, this necessitates reaching down the list of applicants for students who may be less desirable from the admissions standpoint of the particular school (1).

There has been some recent interest in the predictive value of the Strong Vocational Interest Blank in selecting medical students (7). It is well known that there is a high loading on the science factor among successful physicians. The observed tendency in the sub-samples of this study for science grades to be increasingly higher relative to grades in the humanities, as one proceeds from the lowest to highest ranking students of the sub-sample, would seem to be one more indication of the importance of interest configurations emerging from patterns of ability and premedical achievement.

Throughout all segments of education, from the primary grades to the graduate level, there has been a steady growth of concern for student adjustment, especially during the last few decades. As noted above, in approaching the problem of student adjustment and student failures at the Cornell University Medical College, we had come to believe in advance of this study that non-intellectual characteristics of medical students would provide the answer to the student in difficulty, and would be crucial in the case of student failures. The findings that test indications of presence or absence of psychopathology were not predictive, that *MCAT* scores were predictive to a degree, and that more than half of the drop-outs in three classes were recorded as being for non-psychiatric reasons led to the conclusion that preparation and ability are of first importance in success in, and adjustment to, medical school. Once a student is above a certain optimal level of ability and achievement, then the factors of motivation and emotional disturbance will play an increasingly important rôle. But even here, assessment of personality characteristics cannot be used to predict with accuracy future adjustment in the able student.

E. SUMMARY

Three successive classes of entering medical students were tested with the *MMPI* after admission and during the first few weeks of attendance at Cornell University Medical College.

The hypothesis that students who would experience problems in adjustment or psychiatric illnesses during medical training could be located by a quantitative personality inventory was not borne out.

The presence of test indications of clinical levels of psychopathologic functioning was not significantly related to success or failure in medical school. About one quarter of entering students gave such test indications.

More students from these three classes withdrew from medical school for intellectual reasons such as inadequate preparation and an unsuitable ability configuration than for psychiatric reasons.

The best predictors of success in medical school education from admissions data on these students were the *MCAT* total score and the *MCAT* science score which showed a low but significant relationship.

REFERENCES

1. CARLSON, E. T. (Personal Communication.) Cornell University Medical College.
2. CUADRA, C. A. A psychometric investigation of control factors in psychological adjustment. Ph.D. dissertation, Univ. California, 1953.
3. FUNKENSTEIN, D. H. Possible contributions of psychological testing of the nonintellectual characteristics of applicants to medical school. *J. Med. Educ.,* 1957, **32**, 88 (Oct., Part 2).
4. GEE, H. H., & COWLES, T., *Eds.* The appraisal of applicants to medical schools. *J. Med. Educ.,* 1957, **32**, Oct., Part 2.
5. GOTTHEIL, E., & MICHAEL, C. M. Predictor variables employed in research on the selection of medical students. *J. Med. Educ.,* 1957, **32**, 131 (Oct., Part 2).
6. GOUGH, H. B. A new dimension of status: I. Development of a personality scale. *Amer. Soc. Rev.,* 1948, **13**, 401.
7. KELLY, E. L. Multiple criteria of medical education and their implications for selection. *J. Med. Educ.,* 1957, **32**, 185 (Oct., Part 2).
8. KOHL, R. N. The psychiatrist as an advisor and therapist to medical students. *Amer. J. Psychiat.,* 1951, **108**, 198.
9. SCHOFIELD, W. A study of medical students with the *MMPI*: III. Personality and academic success. *J. Appl. Psychol.,* 1953, **37**, 47.
10. WELSH, G. S. An extension of Hathaway's *MMPI* profile coding system. *J. Consult. Psychol.,* 1948, **12**, 343.
11. WOLFLE, D. Medicine's share in America's student resources. *J. Med. Educ.,* 1957, **32**, 10 (Oct., Part 2).

Payne Whitney Psychiatric Clinic
The New York Hospital
525 East 68th Street
New York 21, New York

Published as a separate and in *The Journal of Psychology*, 1959, **47**, 305-316.

VISUAL FIELD AND SUBJECTIVE CENTER OF A DIAMOND*

Department of Psychology, University of Stockholm

THEODOR M. KÜNNAPAS[1]

A. INTRODUCTION AND PROBLEM

The influence of the shape of our monocular visual field on the bisection of a horizontal line was investigated in a previous paper (11). We found that the interocular differences experimentally obtained in the natural visual field were in agreement with the shape of the visual field and that these differences decreased in the circular artificial visual field.

In a simple geometrical figure, the rectangular diamond which is really a square but seems to have a different shape on account of its spatial position, we have two diagonal distances of which one is horizontal and the other vertical. The center of the diamond is equidistant from all four angle points and is therefore the bisection point of both diagonal distances. This quality of the diamond permits the bisection of the horizontal and vertical distances at the same time. In some degree the simultaneous bisection of two distances may facilitate the bisection of one of these. In the bisection of the horizontal diagonal the upper and the lower angle points of the diamond (Figure 1, *c* and *d*) mark approximately the bisection point, and in bisection of the vertical diagonal again the left and right points (Figure 1, *a* and *b*) give some aid. In this meaning the square would be the better figure, but there the diagonal distances are not horizontal and vertical directions. In the diamond, on the other hand, with the increase of the length of the diagonal distances it would be more difficult to use the remaining angle points in the bisection of one diagonal.

In the present study the diamond experiment was undertaken to investigate whether and to what extent the shape of our monocular visual field influences the interocular differences both for the horizontal and the vertical diagonal distances, which may be found in the bisection of these distances by the center point.

Because of the shape of our monocular visual fields and the position of the fixation point in these fields interocular differences between the left and the

---

*Received in the Editorial Office on February 2, 1959, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This investigation was supported by a grant from Magnus Bergvall's Foundation.

right eye are to be expected in bisection of the horizontal diagonal distance. In both monocular visual fields the fixation point is nearer to the nasal boundary than to the temporal. When fixing the center of the diamond with the left eye, the right angle point (b_L) is nearer to the nasal boundary of the visual field than the left point (a_L) to the temporal boundary (Figure 1, A). According to a previous study by the present writer (10) it may be assumed that the right half appears subjectively longer as compared with the left half. In the bisection of this distance into two subjectively equal parts, the left part would be underestimated, and in order to appear as long as the right it should be objectively longer than the right. When we use the right eye, on the contrary, the right half would be underestimated.

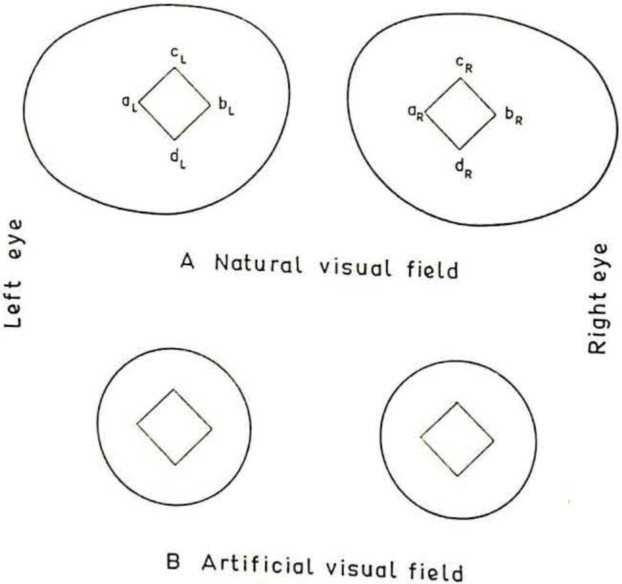


FIGURE 1
THE POSITION OF THE RECTANGULAR DIAMOND IN THE MONOCULAR VISUAL FIELDS: (*A*) THE NATURAL VISUAL FIELDS. (*B*) THE CIRCULAR ARTIFICIAL VISUAL FIELDS

In the case of the experimental artificial visual field both eyes have the same circular shape of the field where the fixation point, as the center of the circle, is in the center. Both in the left and in the right artificial field the distances from the nasal and temporal boundaries to the nearest horizontal angle points of the diamond are all equal (Figure 1, B). Thus no interocular differences like those in the natural monocular fields are to be expected.

However, the results of some previous studies (11, 12) indicated that beside the shape of the visual field, some other factors are of influence. So one may expect that the interocular differences will not disappear totally, but only be reduced in circular artificial visual field.

The vertical diagonal distance has the same position in the natural visual field of both eyes. The distance from the upper angle point (c_L) to the upper boundary of the left visual field is equal to the corresponding distance in the right monocular field (Figure 1, A). Similarly the distances from the lower angle point to the lower boundary are equal in both monocular visual fields. The equal position of the vertical distance does not give rise to any interocular differences in the bisection of that distance. Nor in the circular artificial visual field can these interocular differences be expected to occur.

On the basis of these considerations four quantitatively testable predictions can be made.

1. In the natural visual field the left half is underestimated by the left eye and the right half is underestimated by the right eye, or in the case of the underestimation of the left side even by the right eye, underestimation by the left eye is significantly greater than that by the right eye.
2. In the circular artificial visual field the interocular differences between the left and the right halves decrease significantly.
3. When the vertical diagonal distance is bisected in the natural monocular field, no significant interocular differences are to be expected.
4. The circular artificial visual field does not change the results of the natural visual field in the case of the vertical diagonal.

B. METHOD

1. *Stimulus Pattern*

As stimulus pattern was used a rectangular diamond which had equal horizontal and vertical diagonals. The length of diagonals was 10 cm. The diamond was drawn in the middle of a piece of white paper. The diagonals themselves were not drawn. The lines were 0.5 mm thick.

2. *Apparatus*

In the case of the artificial visual field, the normal monocular visual field was limited by special spectacles. In these spectacles the glasses were replaced by aluminium shields which had the same circular opening. The diameter of the opening was 12 mm. One opening was always covered. The subject could see either with the left or the right eye. Details have been given in a previous paper (12).

3. *Procedure*

The diamond was placed in the middle of a very large white screen on the table. On the screen the contour-lines of the paper on which the diamond was drawn were observed so faintly, especially when the center of the diamond was fixed, that they had no influence as frames. The subject stood in front of the screen, bending over the diamond so that the face was parallel to the table and the center of the figure was between the subject's eyes. The distance from *S*'s eyes to the stimulus pattern was 250 mm. With his right hand *S* held a pencil so that all four angle points of the diamond were visible to him. That pencil came into the surface of the diamond halfway between the lower and the right angle points, and was used by *S* to mark the exact center of the diamond.

The task of *S* was to fix the center of the diamond, to try to find by means of the point of the pencil exactly the point which bisected simultaneously the horizontal and the vertical diagonal distances, and to mark this point on the paper. When fixing the point, *S* had to compare the left part with the right and the upper side with the lower, so that the left part would appear equal to the right, as well as the upper half to the lower. If *S* was not satisfied with the marked center, he had a chance to repeat the trial with another diamond. This possibility was taken advantage of in only some few cases. After the center had been marked, the diamond was taken away by the *E* and a new diamond was placed in the same place on the screen. The average time for *S*s to make a trial was 4 sec. The time between two trials was kept constant at 30 sec.

S viewed monocularly either with his left or with his right eye.

*S*s were made familiar with the experimental situation in preliminary trials.

4. *Experimental Conditions*

Four conditions were conducted:

1. *N_L–Condition.* The visual field was the natural, and *S* used his left eye.

2. *N_R–Condition.* In the case of the natural visual field *S* used his right eye.

3. *A_L–Condition.* In this condition the natural visual field of *S* was limited artificially by spectacles, and he used the left eye.

4. *A_R–Condition.* *S* had the same circular artificial field, and he used his right eye.

The purpose of the Conditions N_L and N_R was to investigate the interocular differences in the natural visual field and to test Predictions 1 and 3. Conditions A_L and A_R had the task of exploring how the natural interocular differences can be influenced by the circular artificial field, and of examining Predictions 2 and 4. Since the position of the head and the distance from the figure were the same in the natural and artificial visual fields and *S* had to fix the same point, so it is to be expected that the same part of the retina would be excited in both cases. Thus, the decrease of the interocular differences in the artificial field in the case of the horizontal diagonal distance is due to the changed shape of the field and not to the change of the excited parts of the retina.

The sequence of the conditions was N_L A_R A_L N_R — N_R A_L A_R N_L or A_R N_L N_R A_L — A_L N_R N_L A_R. The odd numbers of *S*s had the first sequence, the others the second.

5. *Subjects*

Ten university students of psychology, four women and six men, took part in the experiments. Nine *S*s had normal visual acuity, one *S* was slightly myopic (2 dioptries). At the distance of 250 mm the myopic *S* could see the diamond without spectacles as clearly as with them. Each *S* made 8 trials in each of four conditions: four in the first half and four in the other half. Thus, each *S* made 36 trials in all conditions, and the whole group made a total of 360 trials.

C. Results and Discussion

1. *Horizontal Diagonal*

The results for each *S* and for each condition are given in Table 1. The data represent the average differences between the left and right halves when they were equal in apparent length. A positive difference indicates that the left half was longer than the right at the point of subjective equality, and the distances on the left were underestimated relative to the distances on the right. Negative differences, of course, indicate the reverse: the right half was longer and the distances on the right were underestimated relative to distances on the left.

The first two columns show the results in the natural visual field. These agree with the majority of the earlier investigations (1, 3, 5, 6, 7, 8, 9, 11, 15, 16). The point of subjective equality in bisection of the horizontal distances is different for the left and for the right eyes. The left eye underestimates the left side considerably. In case of the right eye the right side is underestimated. The extent of the underestimation is not equal: the left side is underestimated by the left eye much more than the right side by the

TABLE 1

AVERAGE DIFFERENCES IN MILLIMETERS BETWEEN THE LEFT AND RIGHT HALVES WHEN THEY APPEARED EQUAL SUBJECTIVELY IN THE NATURAL ($\mathrm{N_L}$ AND $\mathrm{N_R}$) AND ARTIFICIAL ($\mathrm{A_L}$ AND $\mathrm{A_R}$) VISUAL FIELDS

	Condition			
Subject	$\mathrm{N_L}$	$\mathrm{N_R}$	$\mathrm{A_L}$	$\mathrm{A_R}$
1	1.50	—1.20	.85	— .65
2	2.35	—1.20	1.40	— .55
3	1.10	— .35	.90	— .20
4	1.40	— .90	.60	— .40
5	1.20	.70	.75	1.00
6	3.90	.25	2.10	— .05
7	.50	—1.50	— .55	—1.05
8	.20	—2.45	— .80	— .60
9	.25	—2.15	.35	— .90
10	2.55	1.55	1.20	1.45
All *S*s	1.50	— .73	.68	— .20

right eye. In both cases the temporal parts are underestimated and the nasal halves are overestimated. This resembles the Kundt type of monocular asymmetry and is quite in accordance with Prediction 1.

In the last two columns of Table 1 are given the results for the circular artificial visual field. The point of subjective equality is still different for the left and the right eye. However, the interocular differences, as expected, had decreased considerably. In the natural visual field the interocular difference was 2.22 mm, in the artificial field it was only 0.88 mm. This remarkable decrease was due to the two trends: (*a*) The left eye difference between the left and right halves decreased from 1.50 mm to 0.68 mm, and (*b*) the right eye difference increased from —0.73 mm to —0.20 mm. This result is quite in the line of Prediction 3.

In Figure 2 the left eye is compared with the right eye both in the natural and in the artificial visual fields. The first two columns of the histogram show the primary interocular difference in the natural field. The last two columns indicate the same difference in the artificial field. The comparison of the $\mathrm{N_L}$ column with the $\mathrm{A_L}$, and both (*R*) columns with each other shows to what extent both trends contribute to decrease the interocular difference in the artificial circular field. The rôle of the left eye (*L*) is greater than that of the right eye (*R*).

The differences between the conditions for each *S* are given in Table 2. This table also indicates which differences were significantly different from zero as determined by *t* tests. The difference between the $\mathrm{N_L}$ and $\mathrm{N_R}$ conditions in the natural visual field is significant beyond the .001 level not only for the whole group but also for half of the *S*s. Some *S*s have a lower significance, and only one *S* has no significant difference between the left and right eye. Thus, the left and right eyes have their own points of subjective equality.

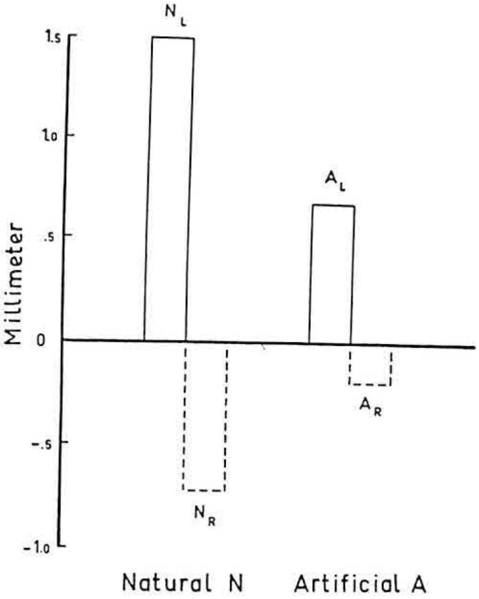


FIGURE 2
THE COMPARISON OF THE LEFT AND RIGHT EYES IN THE BISECTION OF THE HORIZONTAL DIAGONAL IN THE NATURAL AND IN THE ARTIFICIAL VISUAL FIELDS

TABLE 2
DIFFERENCES IN MILLIMETERS BETWEEN CONDITIONS AND THEIR SIGNIFICANCE FOR HORIZONTAL DIAGONAL

Subject	$N_L - N_R$	$N_L - A_L$	$N_R - A_R$	$A_L - A_R$
1	2.70***	.65	— .55	1.50*
2	3.55***	.95	— .65	1.95**
3	1.45**	.20	— .15	1.10*
4	2.30***	.80*	— .50	1.00*
5	.50	.45	— .30	— .25
6	3.65***	1.80**	.30	2.15**
7	2.00**	1.05	— .45	.50
8	2.65***	1.00	—1.85**	— .20
9	2.40**	— .10	—1.25	1.25
10	1.00*	1.35*	.10	— .25
All *S*s	2.22***	.82**	— .53*	.88*

***$P < .001$.
**$P < .01$.
*$P < .05$.

From two trends which had contributed to the decrease of the interocular differences in the artificial field was the trend of the left eye beyond the .01 level and the trend of the right eye beyond the .05 level of confidence. The difference between the left and the right eye in the artificial field remained significant beyond the .05 level, as it was also in case of the bisection of the horizontal line (11).

2. *Vertical Diagonal*

The results for each *S* and for each condition are given in Table 3. The data represent the average differences in millimeters between the upper and the lower halves when they were equal in apparent length. A positive difference indicates that the lower distance was longer than the upper at the point of subjective equality, and distances below were underestimated relative to distances above. Negative differences, of course, indicate the reverse: the upper distances were longer and the distances above were underestimated relative to distances below.

In the natural visual field the point of subjective equality in the bisection of the vertical distances, as opposed to the horizontal distances, is not different for the left and for the right eyes. The left eye has a weak positive difference which means that the lower half is slightly underestimated. Also the right eye underestimated very slightly the same half. So the difference between both monocular visual fields in the bisection of the vertical distance is very small. However, some few *S*s may have greater differences.

In the artificial circular visual field the values of the natural field do not change.

TABLE 3

AVERAGE DIFFERENCES IN MILLIMETERS BETWEEN THE UPPER AND LOWER HALVES WHEN THEY APPEARED EQUAL SUBJECTIVELY IN THE NATURAL (N_L AND N_R) AND ARTIFICIAL (A_L AND A_R) VISUAL FIELDS

	Condition			
Subject	N_L	N_R	A_L	A_R
1	.60	— .15	1.10	.80
2	.65	.60	.50	— .35
3	— .65	— .90	— .70	— .60
4	— .45	— .80	— .25	— .85
5	.35	— .30	— .40	— .05
6	1.95	1.10	1.70	1.45
7	.25	—1.75	—1.40	— .80
8	— .10	— .05	—1.15	—1.05
9	2.45	3.25	2.80	3.00
10	— .35	.75	.70	— .20
All *S*s	.47	.18	.29	.14

The values of the left eye are compared with those of the right eye in Figure 3 both in the natural and the artificial fields. There are no differences worth mentioning between eyes and fields. It is interesting to compare these values with those of the horizontal diagonal in Figure 2.

The exact differences between conditions for the whole group as well as for each *S* and their significance are given in Table 4. These differences are so small that none of them are significant for the whole group, and only few *S*s have some significance in a few cases.

The present results for the natural visual field agree with most previous investigations over the bisection of the vertical line. Delboeuf (2), Mellinghoff (13), Feilchenfeld (3), and Tschermak (15, 16) found also that the lower part in the visual field (or retinally the upper) was underestimated. R. Fischer (4), however, who made experiments with himself and Renqvist (14), reported the underestimation of the upper half. In our group of *S*s, too,

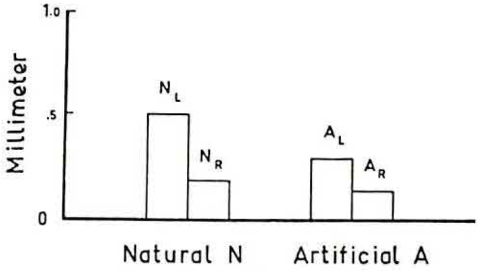


FIGURE 3
THE COMPARISON OF THE LEFT AND RIGHT EYES IN THE BISECTION OF THE VERTICAL DIAGONAL IN THE NATURAL AND IN THE ARTIFICIAL VISUAL FIELDS

TABLE 4
DIFFERENCES IN MILLIMETERS BETWEEN CONDITIONS AND THEIR SIGNIFICANCE FOR VERTICAL DIAGONAL

Subject	$N_L - N_R$	$N_L - A_L$	$N_R - A_R$	$A_L - A_R$
1	.75	— .50	—.95	.30
2	.05	.15	.95	.85
3	.25	.05	— .30	— .10
4	.35	— .20	.05	.60
5	.65	.75	— .25	— .35
6	.85*	.25	— .35	.25
7	2.00***	1.65**	— .95*	— .60
8	— .05	1.05	1.00	— .10
9	— .80*	— .35	.25	— .20
10	—1.10	—1.05*	.95	.90
All *S*s	.30	.18	.04	.16

***$P < .001$.
**$P < .01$.
*$P < .05$.

there were some individuals who had the same underestimation as Fischer, although the group as a whole underestimated the lower half. Brown (1) reported the same interindividual differences, too. Three of his six *S*s underestimated the lower part and the others the upper half. The underestimation of the upper half was still greater. Contrary to our results and to those of the majority of previous investigators, Brown found significant interocular differences.

The subjective O points for each condition are given in Figure 4. The subjective O points of the natural visual field (N_L and N_R) are more remote from the objective O point or from origo than the points of the artificial field (A_L and A_R). The points of the left eye (N_L and A_L) are again more distant from the origo than those of the right eye. Most remote from the objective O point is the subjective O point of the left eye in the natural visual field (N_L), and nearest to the origo is the subjective O point of the right eye in the artificial field (A_R).

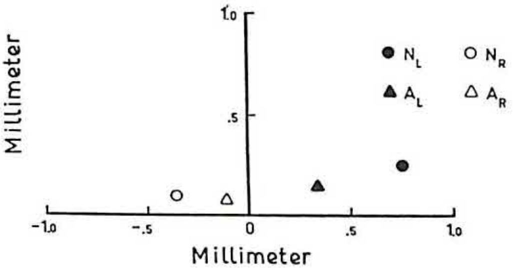


FIGURE 4

The Subjective O Points of All Four Conditions: (*a*) The Left Eye in the Natural Visual Field (N_L); (*b*) the Right Eye in the Natural Visual Field (N_R); (*c*) the Left Eye in the Artificial Field (A_L); (*d*) the Right Eye in the Artificial Field (A_R); the Objective O Point Is O (Origo)

Comparing the results of the vertical diagonal distances with those of the horizontal distance one may say that they were different in the natural visual field. There was a considerable difference in the bisection of the horizontal diagonal between the left and the right eye, corresponding to the different position of this diagonal in relation to the shape of the monocular fields. In the case of the vertical diagonal there was no difference between the position of that diagonal in relation to the shape of the visual field, and also no significant difference between the bisection by the left and right eye.

In the artificial field there were no differences, not only between the position of the vertical diagonal in the left and in the right monocular fields but also between the position of the horizontal diagonal in different monocular visual fields. Corresponding to this condition the results in both cases were very similar to each other, too.

The above considerations support the suggestion that the bisectional differences which are found between the left and right eye are due to the different position of the diagonal distances in relation to the oval shape of the visual field with the longer horizontal axis, and that there is a functional relation between them. We may conclude that all four predictions have been proved to be correct. However, the remaining significance of the interocular differences in the artificial field in the case of the horizontal diagonal, and great interindividual differences indicate that the bisectional difference between the left and right eyes is due not only to the shape of the visual field but also to some other factors.

D. SUMMARY

An attempt is made to investigate whether and to what extent the shape of our monocular visual field influences the interocular differences as well for the horizontal as for the vertical diagonal distances which may be found in the bisection of these distances by the center point of the diamond.

Four conditions of the experimental situation were investigated: (*a*) the left eye in the natural visual field, (*b*) the right eye in the natural visual field, (*c*) the left eye in the artificial visual field, and (*d*) the right eye in the same circular artificial field. In Conditions (*c*) and (*d*) the normal oval visual field was artificially limited by special spectacles with circular openings for the left and right eyes.

In the natural visual field there were found considerable interocular differences in bisection of the horizontal diagonal distance of the diamond, corresponding to the different position of this diagonal in relation to the shape of the monocular visual fields. In the case of the vertical diagonal, where the position of the diagonal was the same in relation to the shape of both monocular fields, the interocular differences in bisection were not significant.

The circular shape of the artificial visual field reduced significantly the interocular differences in the bisection of the horizontal distance, more in the left eye and less in the right eye. In the bisection of the vertical diagonal distance the interocular differences remained very small.

The remaining interocular difference in the artificial visual field in the case of the horizontal diagonal distance, and great interindividual differences indicate the influence of other factors beside the shape of the visual field.

References

1. Brown, K. T. Factors affecting differences in apparent size between opposite halves of a visual meridian. *J. Opt. Soc. Amer., 1953,* **43**, *464-472.*
2. Delboeuf, U. J. *Note sur certaines illusions d'optique: Essai d'une théorie psychophysique de la manière dont l'oeil apprécie les distances, les angles et* les grandeurs. *Bull. l'Acad. Roy. Belgique,* 1865, **19**, 195-216; 1865, **20**, 70-97.
3. Feilchenfeld, H. Über die Grössenschätzung im Sehfeld. *Graefe's Arch. Ophthalm.*, 1902, **53**, 401-422.
4. Fischer, R. Grössenschätzungen im Gesichtsfeld. *Graefe's Arch. Ophthalm.,* 1891, **37**, 97-136.
5. Frank, M. Beobachtungen betreffs der Übereinstimmung der Hering-Hillebrand'schen Horopterabweichung und des Kundt'schen Teilungsversuches. *Pflügers Arch.*, 1905, **109**, 63-77.
6. Herzau, W. Über den Horopter bei schiefer Betrachtung. *Graefe's Arch. Ophthalm.,* 1929, **121**, 757-780.
7. Hillebrand, F. Die Stabilität der Raumwerte auf der Netzhaut. *Z. Psychol.,* 1893, **5**, 1-60.
8. Hofmann, F. B. Physiologische Optik (Raumsinn). *In*: Graefe, A., & Saemisch, T., *Handbuch der gesamten Augenheilkunde.* Berlin: Springer, 1925. (Vol. 3, Chap. 13.)
9. Kundt, A. Untersuchungen über Augenmass und optische Täuschungen. *Annal. Phys. Chem.,* 1863, **120**, 118-158.
10. Künnapas, T. M. The vertical-horizontal illusion and the visual field. *J. Exp. Psychol.,* 1957, **53**, 405-407.
11. ———. Visual field and interocular differences in the bisection of a line. *Acta Psychol.,* 1958, **14**, 375-383.
12. ———. The vertical-horizontal illusion in artificial visual fields. *J. of Psychol.,* 1959, **47**, 41-48.
13. Mellinghoff. Cit. in: Wundt, W., *Grundzüge der physiologischen Psychologie,* II. 5.Aufl. 1902, 560.
14. Renqvist, Y. Versuche über Halbierung von horizontalen, vertikalen und schiefen Strecken. *Skand. Arch. Physiol.,* 1922, **42**, 209-225.
15. Tschermak, A. Über die Grundlagen der optischen Lokalisation nach Höhe und Breite. *Ergebnisse der Physiol.,* 1905, **4**, Abt. 2, 517-564.
16. ———. Optischer Raumsinn. *In*: Bethe, A., Bergmann, G. V., Embdeden, G., & Ellinger, A., *Handbuch der normalen und pathologischen Physiologie.* Vol. 12, pt. 2. Berlin: Springer, 1929-31 (Pp. 834-1000.)

Department of Psychology
University of Stockholm
Teknologgatan 8
Stockholm Va, Sweden

Published as a separate and in *The Journal of Psychology*, 1959, **47**, 317-330.

THE INFLUENCE OF AREA ON THE CRITICAL FLICKER-FUSION THRESHOLD*

The Psychiatric Institute, New York

William C. Roehrig[1]

A. Introduction

Despite the enormous number of articles on the critical flicker-fusion threshold (*CFF*) (18), there have been very few experiments concerned with the influence of one of its basic determinants, the area (*A*) of the test-patch. The few studies that have dealt with this matter have varied *A* over only a narrow range, with the largest area illuminating but a small fraction of the entire retina. In view of the fact that *CFF* is being used more and more frequently as a clinical diagnostic tool in the ophthalmological, psychological, and medical clinics (2, 3, 8, 15, 16, 20, 25), it would seem highly desirable to further explore its major determinants. The dearth of studies of the influence of *A* on *CFF* might perhaps stem from the technical difficulties of homogeneously illuminating a large test-patch with intermittent pulses of light, particularly when high test-patch luminance is desired.

In 1930, Granit and Harper (10) studied the effect of area on *CFF* at the fovea and 10° in the periphery, using six test-patches ranging from 0.9° to 5° of visual angle (*va*) in size. They found that (*a*) the increase in frequency (*F*) with intensity (*I*) was greater, the larger the test-patch; (*b*) *Fmax* (the maximum *F* possible with optimal *I*, under a given set of conditions) was higher, the larger the test-patch; (*c*) the increase in *F* with *A* was greater at 10° in the periphery than at the fovea; and (*d*) the increase in *F* with log *A* was linear over the range tested. They expressed the latter relationship in the form $F = c \log A + d$ (Granit-Harper law), where *c* and *d* are constants. Granit and Harper believed, however, that the linear equation was only an approximation, that the true relationship between *F*

---

*Accepted for publication by Carney Landis of the Editorial Board, received in the Editorial Office on February 3, 1959, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This study is based in part on a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Faculty of Pure Science, Columbia University. This investigation was, in part, made possible by funds granted by the Carnegie Corporation of New York and by grant M-872 from the National Institutes of Health, USPHS.

and log *A* was probably sigmoid, tending to become horizontal below *ca* 1° *va* and above *ca* 5° *va*. Piéron (22) was next to attack the problem, using test-patches ranging from 30″ to 0.5° *va*. His results showed only a slight trend toward becoming horizontal below about 4′ *va*. Hecht and Smith (13), with test-patches of 0.3°, 2°, 6°, and 19° *va*, got contradictory results with two observers (Hecht and Smith). Allen (1) varied the area of the test-patch from about 0.2° to 6° *va* and found not one straight line but a series of straight lines with breaks which he attributed to changes in the retinal distribution of rods and cones. Kugelmass and Landis (17) used a series of 36 test-patches ranging from 1.27° to 14.60° *va* and tentatively agreed with the conclusion of Granit and Harper, that the *F-log A* function was linear from 1° to 5° *va*, becoming negatively accelerated for larger test-patches. However, they failed to find evidence confirming Allen's hypothesis of discontinuities in the *F-log A* function associated with retinal histology.

No fewer than 16 investigators have performed experiments designed to determine what portion of the retina yielded the highest *CFF* values; seven found *CFF* to be highest in the fovea, whereas nine found *CFF* to be highest at various points in the periphery. The reason for these apparently discrepant results was indicated by the following three experiments. Granit and Harper's study (10) showed that with test-patches larger than about 2° *va*, *CFF* was higher 10° in the periphery than at the fovea, whereas with smaller test-patches the reverse was true. Creed and Ruch (6) found *CFF* to be higher in the fovea with a 12′ *va* test-patch, and higher at 4-7° in the periphery with 2° and 8° *va* test-patches. In 1942, Hylkema (15) made a systematic perimetric study of *CFF* with test-patches of 0.5°, 1.5°, 3°, and 10° *va*. He found that with test-patches of 1.5° *va* or smaller, *CFF* was highest in the fovea, whereas larger test-patches yielded higher *CFF* values in the periphery, the highest being approximately 45° out from the fovea, temporally.

Hylkema found that at all portions of the retina tested, the larger the test-patch, the higher the *CFF*. The highest *CFF* (75 *cps*) was considerably higher than has been obtained by most investigators. In an effort to go still higher, Hylkema (14) adjusted his apparatus for its maximum luminance (2.2 *log mL*) and test-patch size (30° *va*). With this combination he obtained an *Fmax* of 82 *cps*, which, to the best of the present writer's knowledge, is the highest that has been reported in almost 50 years. In recent times, other investigators who have reported values nearly as high were Ségal, 79 *cps* (24), and Lloyd, 74.83 *cps* (19). In 1863, Fick (7) performed an experiment with parallel lines on a revolving drum and found an *Fmax* of 170 *cps*, however, this dropped to 40 *cps* if an exit slit was used. Cords (5), in 1908, obtained a *CFF* of 157 *cps*, using no exit slit. Grünbaum (11) published curves in 1898 showing improbably high *F* values of over 650 *cps*. The present writer has so far been unable to decipher Grünbaum's description of his apparatus, and can offer no suggestion as to why his *F* values were so many times greater than those obtained by any of the other hundreds of investigators.

In brief, then, the available evidence indicates that as the area of the test-patch is increased, and particularly as more peripheral portions of the retina are illuminated, *CFF* increases. However, it is unknown how much *CFF* may be made to increase, and what the nature of the relationship between *A* and *CFF* is for other than very small areas. Hence, the two experiments of the present study were designed to (1) investigate the effect of area on *CFF* using foveally-fixated test-patches considerably larger than heretofore used, and with diurnal and day-to-day variability eliminated from the measurements, and (2) determine *Fmax* for practiced observers when area and intensity of the test-patch were as great as the available equipment would permit.

B. Apparatus

The light source was provided by a glow modulator tube (Sylvania Type R1131C). This tube was powered by an electronic driver which caused it to deliver square pulses of light of carefully controlled duration and frequency of intermittence. A block diagram and details of the driver may be found in Kugelmass and Landis (17). Frequency of intermittence was calibrated with an events-per-unit-time meter and was accurate to 0.1 *cps*. Flash duration was calibrated with an oscillator and oscilloscope and adjusted until accurate to 10 μ sec. To insure continuously accurate performance, the driver was always monitored by an oscilloscope. The current to the glow modulator tube, adjustable by means of a potentiometer, was held at 10 *ma* and continuously monitored by a milliammeter. The light-dark ratio was maintained at 1/9 by the proper settings of the dials controlling the duration of the flash and the frequency of repetition.

The optical system used to view the stimulus-patch was a Bausch and Lomb binocular microscope (Model TBV-8L). Mounted rigidly on the stage was an "area-template" which consisted of a brass strip in which accurately spaced holes of increasing diameter had been bored. By means of a rack and pinion gear, the strip could be so moved that one hole at a time was presented and each was accurately centered with respect to the visual axis of the microscope. The areas of the test-patches were defined by these holes. Mounted beneath the stage of the microscope and directly below the area-template was the substage condensing lens of the microscope. Below this lens was a calibrated neutral density wedge having a range of 4 log units, which could be moved back and forth by means of a rack and pinion gear. The top of the glow modulator tube was rigidly fixed below the neutral density wedge in such a way as to be centered with respect to the axis of the optical system. The beam of light emitted by the tube travelled through the neutral density wedge, the condensing lens, the hole in the area-template, and into the head of the microscope which was focused on the edge of the hole. The diameters of the holes in the area-template were measured through the microscope by means of an eyepiece micrometer. Their visual angles were calculated taking into consideration the combined constant magnification of the objective and oculars (30X) and the projection distance which was 250 mm. No artificial pupil was needed since the system of stops in the microscope so limited the cone of light that the diameter of the image of the test-patch at the cornea was less than the smallest natural pupillary diameter.

In order to obtain *Fmax* for as large an area as possible, a second apparatus was used which delivered more light, and was basically different in design from the first. The light source for this apparatus was a 100 *CP* Pointolite lamp, operated with *DC* rectified current having a ripple factor of less than 0.2 per cent which is well within the tolerable limit. By means of a lens system the light was brought to a focus at a point limited by a field stop to 0.1 inch in diameter, where it was chopped by a sector disc which had equal open and closed 90 degree sectors. This disc was turned by a split-phase, four-pole, variable frequency synchronous motor (Electric Indicator Co. Type 450) which would "lock in" with the applied frequency which was precisely controlled by an electronic oscillating unit accurate to 0.1 *cps*. Two separate and complete calibrations of the motor speed were undertaken before the experiment was begun, once with a Strobotac, and once with a built-in calibrating unit which also provided for frequent checks during the course of the experiment. After being chopped by the sector disc the light passed through a filter box and onto a frosted glass screen. In front of this screen was mounted the binocular head of the microscope used in the first apparatus. The area of the test-patch was defined by the system of stops in the microscope, i.e., the largest test-patch area it was possible to obtain. Different luminances were obtained by inserting into the filter box Wratten neutral density filters which had been calibrated with a Martens Densitometer.

All experiments were carried out in a sound-deadened, air-conditioned room in which the level of illumination was approximately three foot-candles.

Data were obtained from three practiced observers, *DD, JL,* and *WR* who were 29, 27, and 28 years old, respectively. For Experiment 2, data were obtained from two additional *O*'s, *CL*, and *JC*, who were 58 and 41 years old, respectively.

C. METHOD

1. *Experiment 1*

The purpose of this experiment was to investigate the effect of area on *CFF* over a large range of areas, the range to include as large an area as the apparatus would permit. Since there was no a priori reason to assume, especially with the large range of *A* to be employed, that the relationship would be the same at all *F* or *I* levels, it was thought desirable to obtain *F-log I* contours for each of the test-patch areas employed. The resultant family of curves could then be "cut" at as many *F* (or *I*) levels as desired and the obtained values plotted against log *A*, thus giving rise to other families of curves which should show any change in the log *A-log I* (or *F-log A*) relationship with increase in *F* (or *I*). If the *F-log I* contours for all the test-patches were not determined at one sitting, it is possible that both diurnal and day-to-day variability might tend to obscure the true relationship between *F* and log *A*. However, since all threshold determinations were to be made at a single sitting, it became necessary to consider carefully the problem of fatigue. Since observation of a flickering light is both unpleasant and fatiguing, threshold determinations in both experiments were obtained only by the method of descending order, that is, from fusion to flicker. Granit and Hammond (9) demonstrated that although *CFF* is affected by duration of observation, there is no further increase in *CFF* values for durations greater than 1 sec. It was concluded from the results of a preliminary experiment not to be reported here, that the results of Granit and Hammond also applied over the range of test-patch areas that were employed in this study. Accordingly, a 3 sec test-duration was used because it was less fatiguing than continuous observation, more comfortable than a 1 sec duration, and gave maximal *CFF* values. The number of threshold determinations at each point was two, in accordance with Hecht's finding (12) that two readings gave as reliable results as 10 or more readings with, of course, far less fatigue. And finally, the number of test-patches was limited, since *F-log I* contours were to be gotten for all at a single sitting.

The test-patches used were chosen so that the portions of the retina they stimulated would correspond approximately with the zones of the retina as defined by Polyak (23, pp. 211-218, and 426). However, note the following by Polyak: "It should be kept in mind, however, that even from a purely morphological standpoint the boundaries between the contiguous regions are more or less arbitrary. In the greater part of the retina the arrangement of the cellular and fibrous elements is essentially the same, the abrupt transitions being absent except in a few places, the change from region to region usually being gradual, almost imperceptible" (23, p. 197). The sizes of the test-patches and the zones of the retina which they stimulated are listed in Table 1. The largest area extended into only a small part of the largest zone of the retina, the far periphery, but it was the largest test-patch possible with the apparatus used.

TABLE 1
TEST-PATCH SIZE AND ZONES OF RETINA STIMULATED

Test-patch diameter in degrees va	Zones of retina and number of degrees va subtended by their outer boundaries (Polyak)		
3.4	I	Inner Fovea	5.0
6.9	II	Parafovea	6.7
17.4	III	Perifovea	16.0
23.5	IV	Near Periphery	24.0
39.7	V	Middle Periphery	40.0
49.6	VI	Far Periphery	

Before beginning the observations, *O* rested his face on a "hood" over the oculars of the microscope through which he made all his observations, adjusted the interocular distance to provide optimal binocular fixation, and adapted to the level of constant illumination for about five minutes. The frequency of intermittence of the test-light was set at 25 *cps*, the lowest frequency used, the area-template set for a 3.4° *va* test-patch, the smallest area of the series, and the neutral density wedge adjusted so that the test-light would appear fused. By pressing his key, *O* presented himself with the test-light for 3 sec, and then made a judgment, either "Flicker" or "Steady." The experimenter increased the luminance in steps of approximately 0.02 log units until the judgment was "Flicker." In this manner, two readings of the critical flicker-fusion luminance (I_c) were taken for each of the six test-patch areas. *O* then rested for several minutes and the procedure was repeated at 35 *cps*, and similarly for 45, 55, 65, 75, 80, and 85 *cps*. The smaller test-patches were eliminated one by one from the series when the next highest frequency setting exceeded *Fmax* for that size test-patch for a given observer. The highest frequency was 85 *cps* for the larger test-patches because the apparatus would not provide sufficient luminance for higher frequencies.

2. *Experiment 2*

The main purpose of this experiment was to utilize the extra light afforded by the second apparatus to determine the nature and extent (*Fmax*) of the *F-log I* contour for the 49.6° *va* test-patch at higher luminances than were possible with the first apparatus. The second apparatus employed the conventional episcotister arrangement, and a light-dark ratio of 1/1. Viewing conditions were similar to those used by the majority of *CFF* investigators, namely, observation was continuous, and *I* was fixed and *F* varied. The diameter of the test-patch was 49.6° *va*, the largest the apparatus would permit. Five *O*'s were used in this experiment.

Before beginning the observations, *O* adapted for about five minutes to the lowest luminance level to be used, with the frequency of intermittence set well above fusion. Then the frequency was lowered slowly and continuously until *O* judged "Flicker," whereupon the frequency was raised well above fusion and the procedure repeated until five readings had been obtained. Similarly for each of 11 succeeding luminance levels which increased by steps of approximately 0.3 log units.

D. RESULTS

1. *Experiment 1*

Table 2 presents, separately for the three *O*'s, the means, in log *mL*, of two determinants of I_c for each of the six test-patches at seven frequencies of intermittence. These data are presented graphically in Figure 1 and have been plotted in the conventional manner with *F* on the ordinate and log I_c on the abscissa, even though in this experiment I_c was the dependent variable. It is seen that the larger the test-patch, the lower was I_c at a given frequency, and the less the increase in I_c with *F*. These results are similar to those of Granit and Harper (10) for small areas (5° *va* or less). The *F-log I* contours all appear to be linear. The negative acceleration usually seen at the top of *F-log I* contours was not found for the small test-patches because the frequency settings were fixed far enough apart to obscure the phenomenon (the *F-log I* contours for the larger test-patches were not extended to their upper limits due to the previously mentioned apparatus limitations).

The data from Table 2 have been replotted in Figure 2 with log I_c as a function of log *A*. It is seen that the relationship was essentially linear over the range tested for all *O*'s. However, as frequency increased, the slopes of the curves increased, i.e., log I_c decreased with increase in log *A* at an increasing rate. This may also be seen in Figure 1 in that the *F-log I* contours spread apart as F increased. Also, for a given frequency, the slope of the log I_c-*log A* curve differed slightly from O to O. This too may be noted in Figure 1 in that for a given frequency, the spread of *F-log I* contours appears greatest for *DD,* less for *WR*, and least for *JL*. Thus, with increase in log A, log I_c decreased in a linear fashion, but at a different rate for different O's and different frequencies.

TABLE 2

CFF THRESHOLDS, I_c (log mL), FOR SIX TEST-PATCHES (DIAM. IN DEGREES VA, AND LOG A MM^2 OF TEST-PATCH), EIGHT FREQUENCIES OF INTERMITTENCE (cps), THREE OBSERVERS, *DD*, *JL*, AND *WR*

	Diam. Log *A*	3.4° 2.21	6.9° 2.85	17.4° 3.66	23.5° 3.93	39.7° 4.41	49.6° 4.62
25 *cps*	*DD*	—0.72	—0.78	—1.05	—1.07	—1.10	—1.17
	JL	—0.71	—0.78	—0.94	—1.01	—1.10	—1.16
	WR	—0.96	—1.15	—1.36	—1.32	—1.42	—1.57
35 *cps*	*DD*	0.06	—0.13	—0.38	—0.48	—0.59	—0.71
	JL	—0.30	—0.47	—0.54	—0.60	—0.68	—0.74
	WR	0.05	—0.17	—0.58	—0.75	—0.81	—0.79
45 *cps*	*DD*	0.84	0.68	0.39	0.23	0.08	—0.04
	JL	0.50	0.35	0.08	—0.06	—0.09	—0.21
	WR	0.98	0.63	0.25	0.08	—0.05	0.00
55 *cps*	*DD*	1.90	1.77	1.03	0.73	0.51	0.41
	JL	1.40	1.16	0.75	0.59	0.51	0.39
	WR	1.77	1.36	0.94	0.79	0.51	0.44
65 *cps*	*DD*	—	—	1.89	1.46	1.04	0.90
	JL	—	1.91	1.45	1.29	1.20	1.08
	WR	—	—	1.57	1.47	1.20	1.02
75 *cps*	*DD*	—	—	—	2.20	1.62	1.41
	JL	—	—	2.16	1.91	1.81	1.73
	WR	—	—	—	—	1.63	1.46
80 *cps*	*DD*	—	—	—	—	1.79	1.64
	JL	—	—	—	—	2.14	2.01
	WR	—	—	—	—	—	—
85 *cps*	*DD*	—	—	—	—	—	2.08
	JL	—	—	—	—	—	—
	WR	—	—	—	—	—	2.10

Figure 3 represents an attempt to translate the measurements obtained by Piéron (22), Allen (1), and the present writer into the same terms for purposes of direct comparison. The *F-log A* contours from the present study were obtained by "cutting" the family of *F-log I* contours for *WR* in Figure 1 at luminance values corresponding to those used by Piéron, 24.5 *mL*, and Allen, 5.3 *mL*. The legitimacy of comparing these "derived" *F-log A* contours rests on the assumption that the *CFF* threshold is the same, whether gotten by varying F or by varying I. In view of the intra- and inter-individual variability of *CFF* data, and the very different experimental procedures used by Piéron, Allen, and the present writer, examination of Figure 3 would seem to justify the conclusion that the *F-log A* relationship is linear from —5.34 to 2.18 log *A* mm^2 on retina, a range of 7.52 log units.

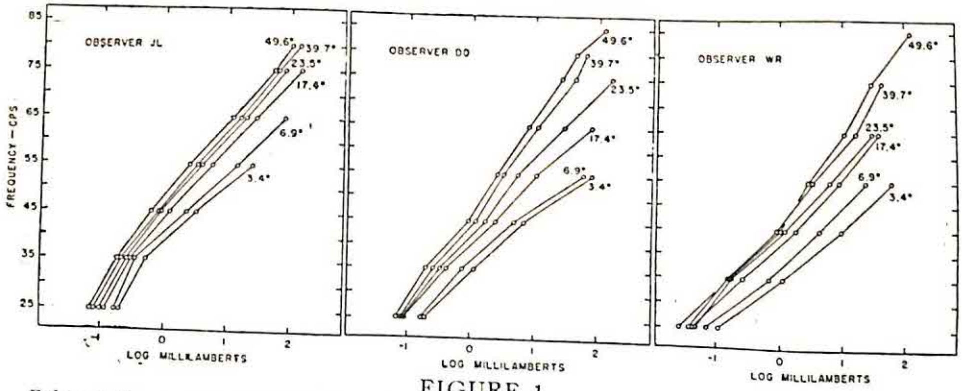


FIGURE 1

F-log I Contours for Six Test-Patches, the Diameters of Which are Indicated to the Right of the Curves in Degrees of Visual Angle

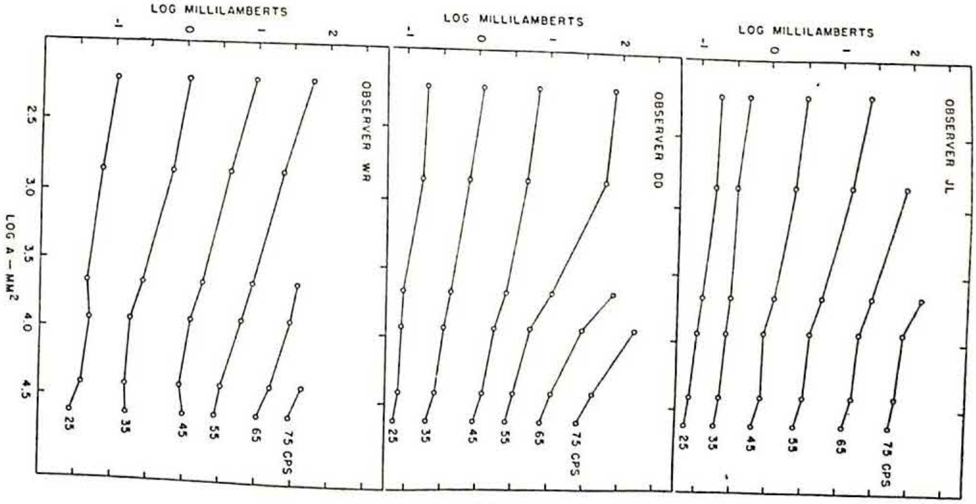


FIGURE 2

Log I_c-Log A Contours for Six Frequency Levels Indicated to the Right of the Curves

2. *Experiment 2*

The results of the second experiment have been presented in Table 3 which lists separately for five *O*'s, the means of five threshold determinations at each of 12 luminance levels. When presented graphically as in Figure 4,

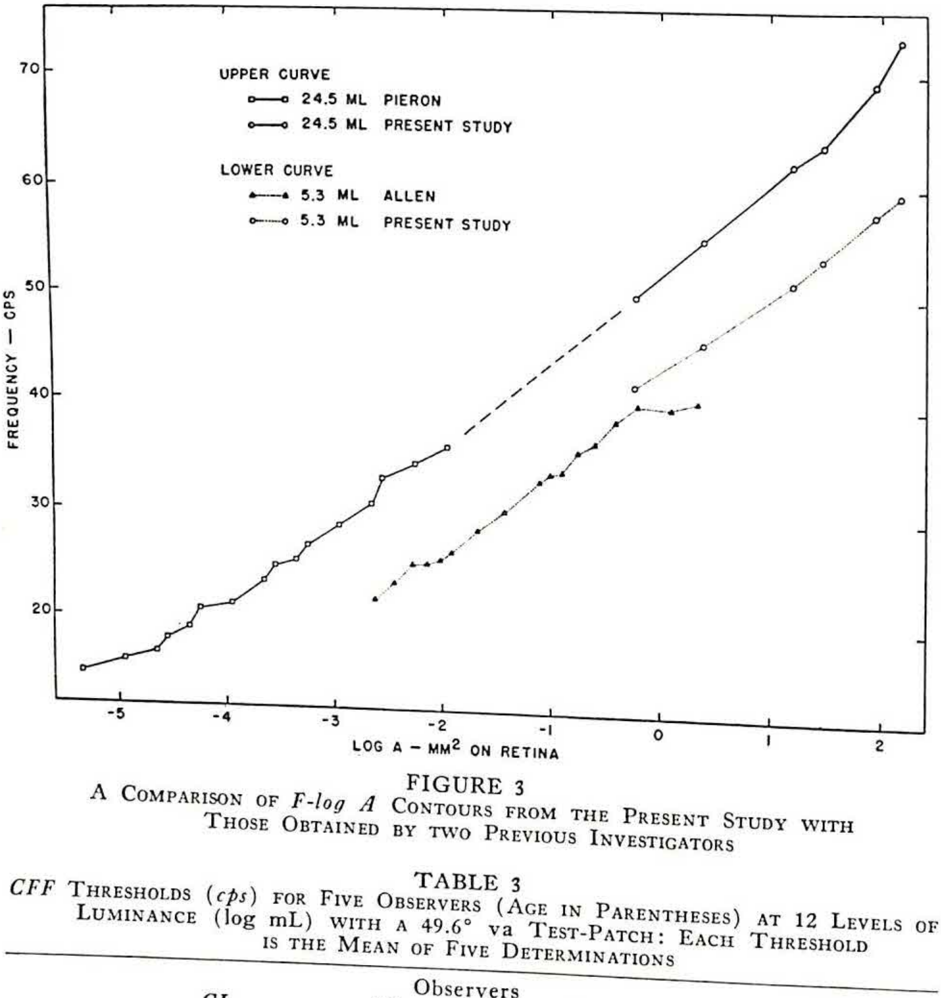


FIGURE 3
A COMPARISON OF *F-log A* CONTOURS FROM THE PRESENT STUDY WITH THOSE OBTAINED BY TWO PREVIOUS INVESTIGATORS

TABLE 3
CFF THRESHOLDS (*cps*) FOR FIVE OBSERVERS (AGE IN PARENTHESES) AT 12 LEVELS OF LUMINANCE (log mL) WITH A 49.6° va TEST-PATCH: EACH THRESHOLD IS THE MEAN OF FIVE DETERMINATIONS

	Observers				
Log mL	*CL* (58)	*JC* (41)	*JL* (27)	*DD* (29)	*WR* (28)
—0.72	24	25	27	28.5	32
—0.37	27	30	33	35	40
0.12	37	40.5	42	43.5	46.5
0.42	44	45	46	49	51.5
0.77	46.5	50	51	53.5	57
1.05	53.5	54.5	56.5	61	65
1.35	58.5	61	62	68	69
1.70	64	66	70	76	77.5
2.12	70	74	78	84	90
2.43	76.5	82	83	90	99
2.64	79.5	86.5	87	95	105
2.77	80	90	89.5	96	107

it may be seen that the *F-log I* contours were all essentially linear over their extent. Although with the added luminance afforded by the second apparatus it was possible to extend considerably the *F-log I* contours for the 49.6° *va* test-patch, the upper limit was still not reached. This will be discussed further in the next section.

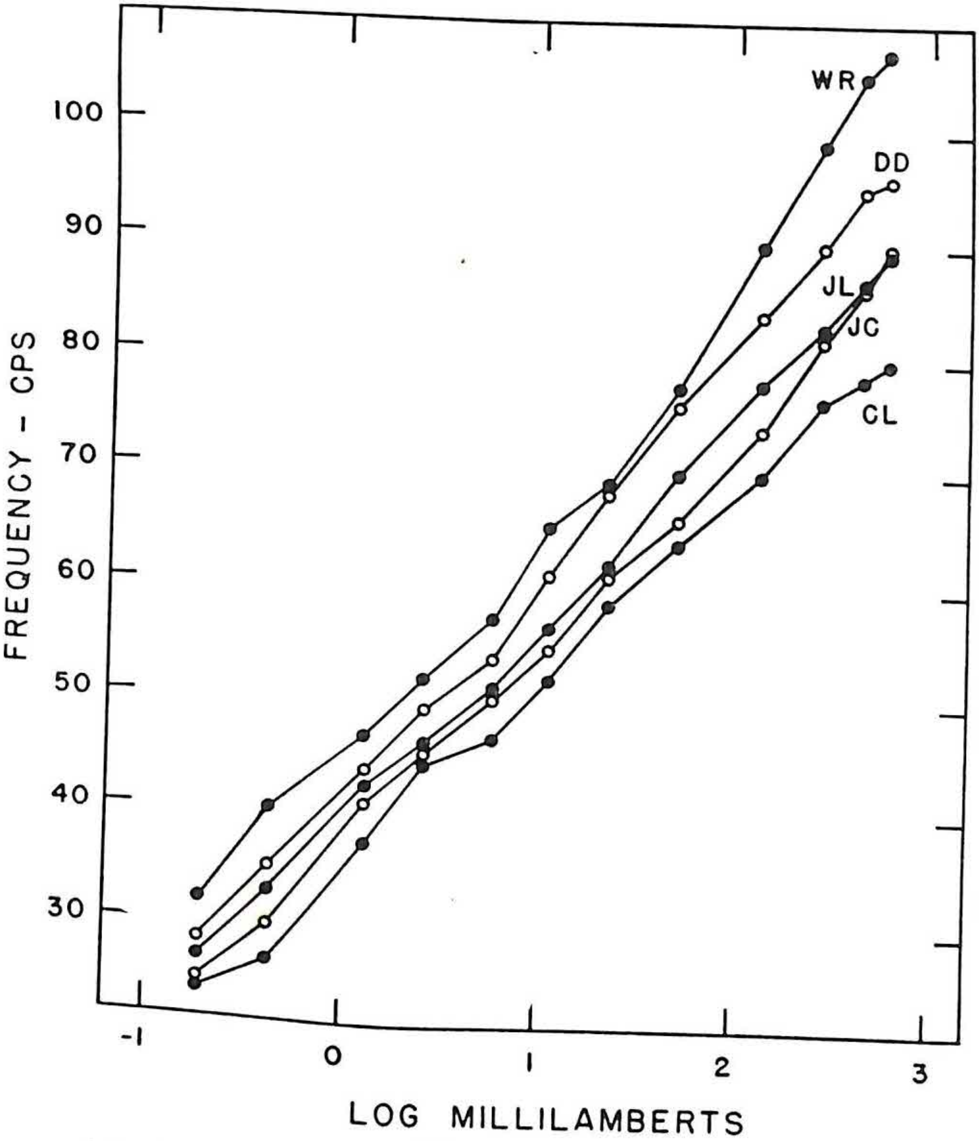


FIGURE 4
F-log I CONTOURS FOR FIVE OBSERVERS WITH A 49.6° VA TEST-PATCH

Two investigators (4, 21) have reported that *CFF* declines with increase in chronological age; however, Misiak found a great deal of individual variability, obtaining *CFF*'s at age 82 as high as at age 7 and *CFF*'s at age 7 as low as at age 80. In the present study, *O*'s *DD, WR,* and *JL* differed but little in age and their *F-log I* contours might have been expected to lie above that of *JC,* whose contour might in turn have been expected to lie above that of *CL.* This was seen to be the case.

E. Discussion

Judging from the fact that the curves in Figures 2 and 3 have not begun to decelerate, it seems clear that the limits of the linear *F-log A* relationship have not yet been reached. It is conceivable that *CFF* would continue to increase with *A* until the whole retina is illuminated. However, since Hylkema's perimetric study (15) showed that for each test-patch area used there was a position on the retina which yielded a maximal *CFF,* one could equally well argue that there is a limit beyond which further increases in *A* would not raise *CFF.* The answer to this question will have to await the surmounting of the technical difficulties alluded to in the introduction. But in any case, the results of the first experiment have demonstrated that the Granit-Harper law covers a much wider range than heretofore believed.

The results of the second experiment show that it is possible to obtain *F* values considerably above the commonly accepted ceiling of 82 *cps*, and since the curves in Figure 4 have not begun to decelerate, it seems reasonable to assume that still higher values might be obtained by further increasing *I.* In fact, temporal resolution by the eye seems to get better and better as *A* and *I* are increased, and the present experiments suggest that *CFF*'s far in excess of 100 may be possible. It should be remembered that with a test-patch of 10° *va,* Hylkema found the maximum *CFF* 45° from the fovea, and the largest test-patch of the present study extended only 25° from the fovea, and covered only a small fraction of the largest zone of the retina, the far periphery.

It is worth noting the linearity of the *F-log I* contours in Figures 1 and 4. As *A* is increased, the *F-log I* contour is extended and its linear portion constitutes an increasingly great proportion. Otherwise stated, the Ferry-Porter law applies to *F-log I* contours for large test-patches as well as small, and describes a greater proportion of the contours for the former.

F. SUMMARY

In the first of two experiments, the effect of area on *CFF* was investigated, using foveally-fixated test-patches considerably larger than heretofore used, and with diurnal and day-to-day variability eliminated from the measurements. When the data were plotted in the form of *F-log I* contours, it was seen that the larger the test-patch, the lower was I_c at a given F, and the less the increase in I_c with F. When the data were plotted with log I_c as a function of log A, it was seen that log I_c decreased linearly with increase in log A, and at an increasing rate with increase in F. At a given F, the rate of decrease in log I_c with increase in log A differed slightly among the three *O*'s; but for all *O*'s the relationship was linear over the range of A investigated. Data from this study were compared with those from two previous investigators and it was concluded that the Granit-Harper law was valid for all test-patch areas so far investigated, i.e., with test-patches from 30″ to 49.6° *va* the relationship between F and log A was linear. It was also suggested that the limit of applicability of the Granit-Harper law had not yet been reached.

It was demonstrated in the second experiment that with a test-patch of 49.6° *va, CFF* values considerably above the commonly accepted top value of 82 *cps* could be gotten, the highest obtained being 107 *cps*. Because the *F-log I* contours had not yet begun to decelerate, it was suggested that had the apparatus been able to deliver larger test-patches and higher luminances, still higher *CFF* values might have been obtained. It was noted that the *F-log I* contours in this study all appeared linear, and hence the Ferry-Porter law applies to *F-log I* contours for large test-patches as well as small.

REFERENCES

1. ALLEN, F. The delineation of retinal zones with dark tube vision. *Canad. J. Res.*, 1945, **23A**, 21-31.
2. BRILL, H. M., GOLDEN, M. J., WOOLF, H. L., KLAWANS, A. H., & LONG, J. S. The preclinical recognition of toxemia of pregnancy. *Amer. J. Obstet. & Gynaec.*, 1951, **62**, 614-620.
3. BRILL, H. M., LONG, J. S., KLAWANS, A. H., GOLDEN, M., & SEAMAN, I. The nitroglycerin flicker fusion test in the toxemia of pregnancy. *Amer. J. Obstet. & Gynaec.*, 1952, **64**, 1201-1210.
4. COPPINGER, N. W. The relationship between critical flicker frequency and chronological age for varying levels of stimulus brightness. (Unpublished Ph.D. dissertation, Tulane University, 1951.)
5. CORDS, R. Über die Verschmelzungsfrequenz bei periodischer Netzhautreizung durch Licht oder elektrische Ströme. *v. Graefes Arch. Ophthal.*, 1908, **67**, 149-161.
6. CREED, R. S., & RUCH, T. C. Regional variations in sensitivity to flicker. *J. Physiol.*, 1932, **74**, 407-423.

7. Fick, A. Ueber den zeitlichen Verlauf der Erregung in der Netzhaut. *Arch. Anat. Physiol.*, 1863, 739-764.

8. Golden, L. H., & Cordasco, E. M. Flicker photometry and electrocardiography: A correlation. *N. Y. S. J. Med.*, 1952, **52**, 2157-2158.

9. Granit, R., & Hammond, E. L. Comparative studies on the peripheral and central retina: V. The sensation-time curve and the time course of the fusion frequency of intermittent stimulation. *Amer. J. Physiol.*, 1931, **98**, 654-663.

10. Granit, R., & Harper, P. Comparative studies on the peripheral and central retina: II. Synaptic reactions in the eye. *Amer. J. Physiol.*, 1930, **95**, 211-228.

11. Grünbaum, O. F. F. On intermittent stimulation of the retina: I. *J. Physiol.*, 1897, **21**, 396-402.

12. Hecht, S., Shlaer, S., & Verrijp, C. D. Intermittent stimulation by light: II. The measurement of critical fusion frequency for the human eye. *J. Gen. Physiol.*, 1933, **17**, 237-249.

13. Hecht, S., & Smith, E. L. Intermittent stimulation by light: VI. Area and the relation between critical frequency and intensity. *J. Gen. Physiol.*, 1936, **19**, 979-991.

14. Hylkema, B. S. Fusion frequency with intermittent light under various circumstances. *Acta Ophthal.*, 1942, **20**, 159-180.

15. ———. Examination of the visual field by determining the fusion frequency. *Acta Ophthal.*, 1942, **20**, 181-193.

16. Krasno, L. R., & Ivy, A. C. The response of the flicker fusion threshold of nitroglycerin and its potential value in the diagnosis, prognosis, and therapy of subclinical cardiovascular disease. *Circulation*, 1950, **1**, 1267-1276.

17. Kugelmass, S., & Landis, C. The relation of area and luminance to the threshold for critical flicker fusion. *Amer. J. Psychol.*, 1955, **68**, 1-19.

18. Landis, C. An Annotated Bibliography of Flicker Fusion Phenomena Covering the Period 1740-1952. Armed Forces National Research Council, Vision Committee Secretariat, University of Michigan, Ann Arbor, Michigan, 1953.

19. Lloyd, V. V. A comparison of critical fusion frequencies for different areas in the fovea and periphery. *Amer. J. Psychol.*, 1952, **65**, 346-357.

20. Miles, P. W. Flicker fusion fields: II. Technique and interpretation. *Amer. J. Ophthal.*, 1950, **33**, 1069-1077.

21. Misiak, H. The decrease of critical flicker frequency with age. *Science*, 1951, **113**, 551-552.

22. Piéron, H. L'influence de la surface rétinenne en jeu dans une excitation lumineuse intermittente sur la valeur des fréquences critiques de papillotement. *C. R. Soc. Biol.*, 1935, **118**, 25-28.

23. Polyak, S. L. The Retina. Chicago: Univ. Chicago Press, 1941.

24. Ségal, J. Le méchanisme de la vision en lumière intermittente. *J. Psychol., Norm. & Path.*, 1939, **36**, 451-539.

25. Young, K. H. Critical flicker frequency. *In* F. A. Mettler (*Ed.*), *Selective Partial Ablation of the Frontal Cortex*. New York: Hoeber, 1949. (Pp. 257-263.)

The Psychiatric Institute
722 West 168th Street
New York 32, New York

Published as a separate and in *The Journal of Psychology*, 1959, **47**, 331-335.

ABILITY VS. PRACTICE IN PAIRED-ASSOCIATE LEARNING*

Department of Psychology, Montana State University

CLYDE E. NOBLE[1]

A. PROBLEM

Recent studies of the relationship between individual differences in ability (A) and amount of practice (N) on such psychomotor tasks as rotary pursuit (10), two-hand coördination (3), and discrimination reaction (1, 7) have found conflicting evidence regarding possible interaction between these two variables. Experiments on more discrete S-R tasks like paired-associate syllable learning (2, Exp. II) and non-verbal trial-and-error learning (9), on the other hand, typically reveal systematic tendencies for Ss' rate of growth to be positively correlated with initial level of ability. This is especially true when the likelihood of positive transfer is minimized or when Ss' performance begins approximately at chance.

The present study reports further data testing the hypothesis that interaction laws of the type $R\% = f\ (A \times N)$ are characteristic of *selective learning* (6), whatever may be the principle governing the acquisition of *perceptual-motor skill*. Performance curves in paired-associate verbal learning will be analyzed for the presence of an intrinsic ability by practice interaction by dividing the sample into several A-levels based on initial proficiency, then plotting the $R\%$ curves of these homogeneous subgroups as a function of N. Jointly under evaluation, it may be noted, is the more inclusive hypothesis of Hull (4) that individual differences can be represented by parametric variations of general behavior equations.[2] Research on this problem is expected to yield information relevant to the controversy between idiographic (particular) and nomothetic (general) modes of description. In one form or another, this issue has been recurring in psychology ever since the Titchener-Baldwin debate of the 1890s.

---

*Received in the Editorial Office on February 18, 1959, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The data were collected by Deldon A. McNeely at Louisiana State University in 1957. Statistical analyses were performed at Montana State University by J. E. Fuchs, Janet L. Noble, and H. A. Walters, supported in part by a grant from the M.S.U. Research Committee. Drs. J. A. Adams and R. B. Ammons made helpful suggestions.

[2] An application of Hull's notion to human trial-and-error learning is reported in (9).

B. METHOD

1. *Apparatus*

The learning device was a Patterson memory drum which presented the paired-associates typed in capital letters at a 4-sec. rate. The *S* term was exposed for 2 sec., followed by *S* and *R* together for 2 sec. The intertrial interval was 4 sec.

All *S*s practiced a difficult list consisting of 10 pairs of relatively meaningless and unfamiliar two-syllable nouns or paralogs. As in a previous experiment (2, Exp. I), the items were drawn from the low end of Noble's (5) scale of meaningfulness (m) such that the mean median m value was 0.29. Pairs were combined to avoid mediated associations and similar sources of bias. There were four different random trial sequences, and the series repeated on Trials 5, 9, 13, and 17.

2. *Subjects*

The *S*s were 70 undergraduate psychology students, 15 women and 55 men, whose ages ranged from 18 to 26 (mean = 21.3 yr.). All were naïve with respect to paired-associate learning.

3. *Procedure*

Each *S* received 20 trials by the anticipation-correction procedure, and both *S* and *R* were pronounced. Instructions were the same as those in prior studies (2, 8), but no preliminary practice was given. Each *S* began with the same trial sequence, and all were treated alike during training. Following the tabulation of the data, *S*s were stratified into four homogeneous ability levels (A). The basis of the A classification was the total number of correct responses (R+) made by *S* during the first eight trials. Each subgroup contained 10 *S*s selected to ensure minimum intragroup variability and approximately equal intervals along the ability continuum, as in (3). The cumulative R+ scores earned by the four subgroups by the end of Trial 8 fell in the following ranges: I (slow) = 0, II (below average) = 2-3, III (above average) = 5-6, and IV (fast) = 8-10. The means were 0, 2.9, 5.5, and 8.7 R+'s respectively.

C. RESULTS AND DISCUSSION

Figure 1 shows the percentage of correct responses ($R\%$) as a function of practice (N) by two-trial blocks, with initial ability (A) as the parameter. There is a tendency toward negative acceleration in all subgroups, but it is less pronounced as ability increases. Inverse relationships between ability and variability have been observed in earlier selective learning experiments (8, 9), so Bartlett's test was applied to the total $R+$ scores of each S in the four ability strata. For 3 df, $\chi^2 = 5.44$ which is not significant ($p > .10$), indicating that the subgroups are not heterogenous in variance. A 4x19 mixed-factorial analysis of variance performed on the $R+$ scores of the four subgroups over Trials 2-20 is summarized in Table 1. Both main effects and the $A \times N$ interaction are significant at or beyond the .1 per cent point, as expected. Figure 1 and Table 1 are therefore consistent with the hypothesis

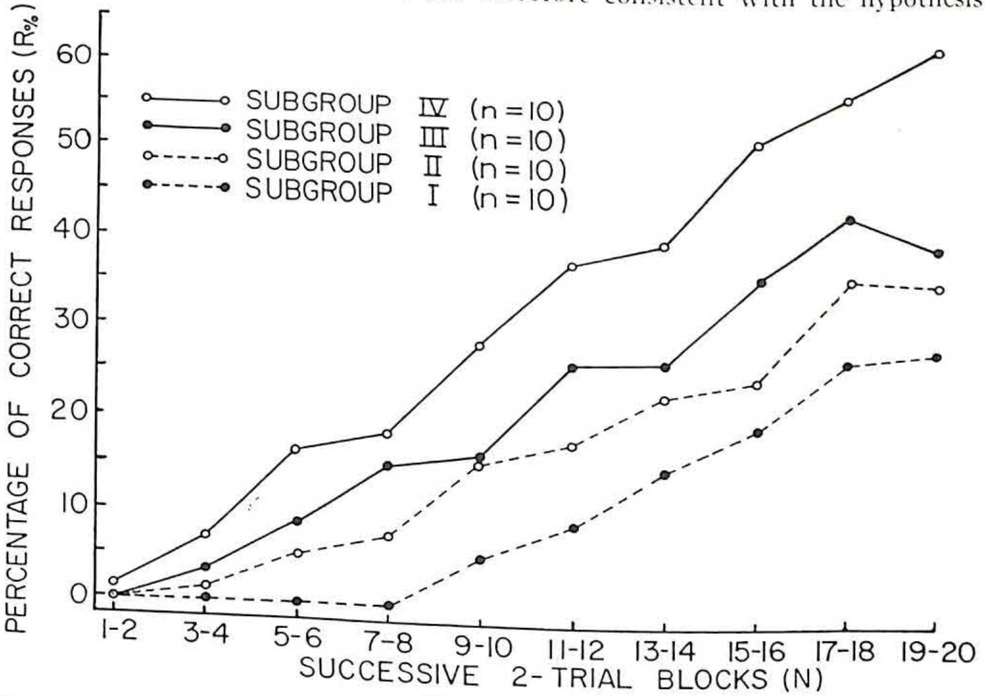


FIGURE 1

Acquisition Curves for Lists of 10 Paired Associates as a Function of Practice With Initial Ability as the Parameter: Each Subgroup Was Selected for Homogeneity on the Basis of Performance During Trials 1-8

TABLE 1

Analysis of Variance of Total Correct Responses ($R+$) During Trials 2-20 for Four Ability Levels

Source	df	MS	F
Between Ss	39		
Ability Level (A)	3	173.63	23.24*
Error (b)	36	7.47	
Within Ss	720		
Trials (N)	18	85.55	90.05*
Interaction ($A \times N$)	54	1.99	2.09*
Error (w)	648	.95	
Total	759		

*$P < .001$.

that rate of acquisition of paired-associate proficiency is a positive function of initial level of ability; this confirms Hull's idea too (4, 9).

As Spence (11) has pointed out, paired-associate learning is similar to classical conditioning in that specific *S-R* connections may be isolated in relative independence of competing response tendencies. When homogeneous groups of human *S*s in eyelid conditioning are formed in the manner of this study, the stratified acquisition curves present the same skewed, sigmoid appearance as those in Figure 1. Another example of selective behavior, compound trial-and-error learning (6, 9), also agrees with the generalization that frequency ($R\%$, R_p) curves in both simple and complex *S-R* learning situations are *S*-shaped, with the degree of skewness and initial positive acceleration contingent upon the ability level of the *S*s.

Why perceptual-motor skill curves do not consistently reveal an $A \times N$ interaction is not clear, but a number of possibilities exist. Among these are: (*a*) large amounts of initial positive transfer from pre-laboratory situations, (*b*) failure to extend the training period until an asymptotic phase is approached, (*c*) excessively coarse grouping of trials, (*d*) lack of comparability of the response criteria being measured, and (*e*) "floor" and/or "ceiling" restrictions distorting the units of measurement. Much work remains to be done before the fields of selective learning and psychomotor skill can be coördinated; meanwhile, the analytic properties of the former (6) appear to lend themselves well to experimental and theoretical treatment in *S-R* terms (9).

D. SUMMARY

Seventy *S*s were given 20 trials of practice on a paired-associate verbal list of low meaningfulness ($m = 0.29$). After dividing the sample into different ability levels based on proficiency during Trials 1-8, the acquisition curves of four homogeneous subgroups of 10 *S*s each were examined for the presence of an ability by practice interaction. Analysis of variance indicated that the effects of practice, initial ability level, and their interaction were all significant. These data confirm previous experiments on classical conditioning and selective learning which show *S*s' acquisition rate to be positively correlated with initial level of ability.

REFERENCES

1. ADAMS, J. A. The relationship between certain measures of ability and the acquisition of a psychomotor criterion response. *J. Gen. Psychol.*, 1957, **56**, 121-134.
2. CIEUTAT, V. J., STOCKWELL, F. E., & NOBLE, C. E. The interaction of ability and amount of practice with stimulus and response meaningfulness (m, m') in paired-associate learning. *J. Exp. Psychol.*, 1958, **56**, 193-202.

3. Cieutat, V. J., & Noble, C. E. Ability vs. practice in two-hand coördination. *Percept. & Mot. Skills*, 1958, **8**, 226.
4. Hull, C. L. The place of innate individual and species differences in a natural-science theory of behavior. *Psychol. Rev.*, 1945, **52**, 55-60.
5. Noble, C. E. An analysis of meaning. *Psychol. Rev.*, 1952, **59**, 421-430.
6. ———. Human trial-and-error learning. *Psychol. Rep.*, 1957, **3**, 377-398.
7. Noble, C. E., Fuchs, J. E., Robel, D. P., & Chambers, R. W. Individual vs. social performance on two perceptual-motor tasks. *Percept. & Mot. Skills*, 1958, **8**, 131-134.
8. Noble, C. E., & McNeely, D. A. The rôle of meaningfulness (*m*) in paired-associate verbal learning. *J. Exp. Psychol.*, 1957, **53**, 16-22.
9. Noble, C. E., Noble, J. L., & Alcock, W. T. Prediction of individual differences in human trial-and-error learning. *Percept. & Mot. Skills*, 1958, **8**, 151-172.
10. Reynolds, B., & Adams, J. A. Psychomotor performance as a function of initial level of ability. *Amer. J. Psychol.*, 1954, **67**, 268-277.
11. Spence, K. W. Behavior Theory and Conditioning. New Haven: Yale Univ Press, 1956.

Department of Psychology
Montana State University
Missoula, Montana

Published as a separate and in *The Journal of Psychology,* 1959, **47**, 337-341.

A FURTHER STUDY IN THE CONSISTENCY OF JUDGMENTS IN CATEGORIZING VERBAL MATERIAL*

New York University and Bell Telephone Laboratories

M. E. Tresselt and M. S. Mayzner[1]

A. Problem

Previous studies (1, 2, 3, 4, 5) have explored some of the variables which influence the number of words *S*s categorize as belonging to such verbal concepts as theoretical, economic, aesthetic, social, political, and religious. In all of these studies, *S*s were presented with a list of words and were asked to check those they judged belonged to the given concepts. Thus, for a given word *S* had two options, with respect to a given concept, the word belonged or it did not belong. A study (5) designed to examine the consistency of this type of judgment found it to be very high.

The present study basically is concerned with the consistency of judgments in categorizing verbal material, when two different judgmental procedures are compared. Specifically, a single verbal concept, namely economic, was chosen, and the results obtained in a previous study (4) for this concept, employing only two judgmental options (belonging or not belonging) were compared with a procedure in which *S* rated the words as belonging or not belonging to this concept on a seven-point rating scale. In addition, judgments on these words, using the seven-point rating scale, were elicited for one group of *S*s approximately 10 minutes after completion of their first judgments and for a second group of *S*s approximately 48 hours after completion of their first judgments. The foregoing procedure was included in the study to provide information concerning the reliability associated with such verbal judgments.

B. Method

1. *Subjects*

One hundred *S*s obtained from undergraduate psychology courses at New York University were employed in the study.

---

*Received in the Editorial Office on February 19, 1959, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This study was supported in part by a grant made to M. E. Tresselt by the New York University Arts and Science Research Fund.

2. *Materials*

The stimulus material presented to *S*s consisted of a list of 100 words (shown in Table 1). These 100 words were selected from a larger sample of 600 words employed in an initial scaling study (4). In this initial study, 100 *S*s were asked to judge or check each of 600 words as belonging to none, 1, 2, 3, 4, 5, or 6 concepts, the concepts being theoretical, economical, aesthetic, social, political, and religious. By summing the number of checks for each word for the 100 *S*s, 600 word profiles were obtained, which showed the frequency with which each word was judged as belonging to each concept. Using only the scale values (i.e., the number of checks for a given concept) for the economic concept, 100 words were selected for this study, so that the words appeared as nearly as possible at equal intervals along the scale of checkability. That is, the words in Table 1 are arranged in order of increasing checkability (as obtained in the initial scaling study), and the associated scale values (shown under Column I) range from 1 to 100, in equal steps of one unit, except in a few cases, where there is a two- or three-unit jump.

3. *Procedure*

The *S*s were presented with the words in Table 1, completely randomized as to scale value, on mimeographed sheets. They were instructed to judge the degree to which each word belonged to the category economic using a seven-point scale. The number "one" was to be used if they judged the word as not belonging or belonging to a very small degree to the category; the number "seven" was to be used if they judged the word as belonging completely or to a very strong degree to the category. The intermediate numbers, "two" through "six" were to be used to express intervening degrees of belonging. Fifty of the 100 *S*s tested were administered the same list approximately 10 minutes after completion of their first judgments, the remaining 50 *S*s were administered the same list approximately 48 hours after completion of their first judgments, with all lists being administered during regular class periods.

C. RESULTS AND DISCUSSION

The major results of the study are given under Column II in Table 1. The scale values shown were calculated by averaging the ratings (1 through 7) given by the 100 *S*s for each word. The standard deviations associated with the mean scale values are also given. The mean scale values range from 1.17 to 6.67, demonstrating a fairly good spread over the total range possible i.e., 1.00 to 7.00.

Since the basic objective of the study was to determine the consistency

TABLE 1
WORDS AND ASSOCIATED SCALE VALUES USED IN THE STUDY

Words	Scale values *I**	*II**M*	*SD*	Words	Scale values *I*	*II-M*	*SD*
1. Poem	1	1.17	.89	51. Liquidate	51	4.57	2.24
2. Horizon	2	1.86	1.52	52. Partner	52	4.84	1.87
3. Priest	3	1.43	1.19	53. Volume	53	4.61	2.06
4. Kick	4	1.21	.84	54. Union	54	5.70	1.61
5. Blend	5	2.16	1.69	55. Power	55	5.17	1.89
6. Skeptic	6	2.08	2.42	56. Report	56	3.67	1.95
7. Classical	7	1.57	1.19	57. Policy	57	5.14	1.95
8. Loyalty	8	2.34	1.74	58. Quality	58	4.61	2.09
9. Counterpoint	9	2.09	1.67	59. Personnel	59	5.20	1.85
10. King	10	2.68	2.28	60. Save	59	5.06	1.78
11. Banquet	11	2.30	1.79	61. Career	61	4.65	2.21
12. Violent	12	1.73	1.49	62. Ransom	62	2.91	2.34
13. Illumination	13	2.29	1.85	63. System	63	5.43	1.79
14. Investigate	14	3.48	1.84	64. Utilize	64	5.02	1.96
15. Civil	15	3.33	2.05	65. Reward	65	3.43	2.03
16. Appeal	16	2.80	1.99	66. Boss	65	4.99	1.89
17. Execute	17	2.85	1.96	67. Statistics	67	5.41	1.77
18. Evolution	18	3.01	2.12	68. Work	68	6.03	1.37
19. Judgment	19	4.05	2.04	69. Value	69	5.59	1.66
20. Tea	20	1.97	1.72	70. Office	71	4.37	1.77
21. Trunk	21	1.28	1.09	71. Bonds	72	6.42	1.15
22. Comparison	22	3.02	2.02	72. Trust	72	5.66	1.89
23. Chair	23	1.53	1.27	73. Department	73	3.45	1.88
24. Form	24	2.53	1.84	74. Increase	74	4.20	1.92
25. Travel	25	3.27	2.00	75. Client	75	4.54	1.34
26. Method	26	4.16	2.07	76. Demand	76	5.40	2.93
27. Selfish	27	2.94	2.11	77. Appropriation	77	5.66	1.70
28. Entry	28	3.28	2.19	78. Competition	78	6.12	1.48
29. Win	29	3.05	2.14	79. Margin	79	5.22	2.01
30. Blanket	30	1.64	1.40	80. Cartel	80	5.54	2.00
31. College	31	3.98	1.78	81. Consumption	82	5.87	1.91
32. Object	32	2.34	1.78	82. Machinery	83	4.95	1.91
33. Drive	33	3.89	2.21	83. Petroleum	83	4.66	2.05
34. Experience	34	3.62	2.02	84. Coal	84	4.07	2.10
35. Free	35	3.46	2.31	85. Rich	85	5.10	2.08
36. Society	36	5.13	1.94	86. Imports	86	6.30	1.11
37. Restriction	37	3.97	2.07	87. Management	87	5.92	1.45
38. Concrete	38	3.14	2.05	88. Rates	89	5.78	1.55
39. Socialism	39	4.56	2.76	89. Receipt	89	5.10	1.73
40. Standards	40	5.33	1.79	90. Budget	92	6.44	1.19
41. Charity	41	3.26	2.06	91. Expenditure	92	5.85	1.50
42. Bridge	42	2.90	1.26	92. Sale	92	5.80	1.45
43. Area	43	2.95	2.10	93. Monopoly	93	6.40	1.28
44. Company	44	5.37	1.82	94. Depreciation	94	5.58	1.69
45. Balance	45	5.17	1.94	95. Invest	95	6.54	.96
46. Flood	47	3.41	2.28	96. Trade	96	6.36	1.00
47. Organize	47	5.46	1.75	97. Assets	96	6.20	1.43
48. Wear	48	2.23	1.79	98. Finance	98	6.67	.85
49. Deal	49	4.62	2.13	99. Wages	99	6.35	1.24
50. News	50	3.60	2.14	100. Industry	100	6.32	1.23

*Column I gives the scale values (frequency of checkability) for the words from the initial scaling study.

**Column II gives the scale values for the words obtained in the present study.

between two different judgmental procedures, a Rank Difference correlation was calculated for the two sets of scale values given in Table 1, under Columns I and II. The value, rho, obtained was .86, which is highly significant ($P < .0001$). In addition, 25 *S*s each were selected at random from the 50 *S*s available in the 10-minute and 48-hour delayed test conditions, and Pearson Product-Moment correlations were calculated for each *S* to determine test-retest reliability of the judgments made. The correlations from the 10-minute delayed test condition ranged from .67 to .91, all highly significant ($P < .0001$). The correlations for the 48-hour delayed test condition ranged from .64 to .88, again all highly significant. These results clearly show that the reliability in making such judgments is consistently high for a given *S* whether he makes his second set of judgments soon after the first (10 minutes) or after a considerable delay (48 hours).

The foregoing analysis has demonstrated that the type of categorization under consideration essentially is independent of the type of specific judgmental technique used to measure it. A previous study (5) showed a high degree of consistency for such categorization by two independent groups of *S*s judging the same verbal stimuli some three years apart. The present study has shown a high correlation between two different judgmental techniques used in measuring such categorization and a high reliability for one technique on replication. The two studies taken together then present a picture of outstanding stability for such verbal categorizing behavior and raises the general question as to how such stability is achieved.

D. SUMMARY

The present study was designed to discover the consistency between two different judgmental techniques employed in categorizing verbal stimuli, and the reliability of one of these techniques. In one technique *S*s were asked to judge whether words did or did not belong to given conceptual categories, and in the second technique *S*s expressed their judgments on a seven-point rating scale. The results clearly demonstrated the high consistency between the two techniques, with a Rank Difference correlation of .86 being obtained, and the high reliability of one of them (the seven-point rating scale), with Pearsonian *r*'s ranging from .64 to .91.

REFERENCES

1. MAYZNER, M. S.. & TRESSELT, M. E. Concept span as a composite function of personal values, anxiety, and rigidity. *J. Pers.*, 1955, **24**, 20-33.
2. ———. Shifts in connotative meaning of words as a function of previous restrictive experience. *J. Exp. Psychol.*, 1958, **55**, 200-205.
3. ———. Shifts in connotative meaning of words as a function of varying amounts of previous restrictive experience. *J. of Psychol.*, 1959, **47**, 107-116.

4. PRITCHARD, L., MAYZNER, M. S., & TRESSELT, M. E. The Mayzner-Tresselt word-concept check list. *Psychol. Newsl.*, 1955, **7**, 27-33.
5. TRESSELT, M. E., & MAYZNER, M. S. Consistency of judgments in categorizing verbal material. *Psychol. Rep.*, 1958, **4**, 415-421.

Department of Psychology
New York University
Washington Square
New York 3, New York

Bell Telephone Laboratories
Whippany Laboratory
Whippany, New Jersey

Published as a separate and in *The Journal of Psychology*, 1959, **47**, 343-352.

WEBER'S LAW AND RELATED FUNCTIONS*

Department of Psychology, University of Stockholm

GÖSTA EKMAN[1]

A. INTRODUCTION AND PROBLEMS

As is well known, Fechner's law was deduced from Weber's law on the assumption that the subjective correlate ($\triangle R$) of a just noticeable stimulus difference ($\triangle S$) is constant over R and thus may be used as an arbitrary unit of measurement.

The invalidity of Fechner's law has been repeatedly demonstrated, and there is very convincing evidence that it must be replaced by a power function. This has been clearly shown, at present for some 20 or more continua, especially by Stevens and his co-workers (16, 17). It has also been shown that $\triangle R$ tends to increase with R, at least for "prothetic" continua (17).

The purposes of the present paper are: (*a*) to investigate in some more detail the relation between R and $\triangle R$ in four subjective continua; (*b*) to derive, on the basis of this experimentally established relation, an adequate form for Weber's law as applied to the stimulus continuum, and to test the function by fitting it to experimental data; and (*c*) to discuss in a preliminary way some possible relations between differential and absolute sensitivity.

B. DIFFERENTIAL SENSITIVITY IN FOUR SUBJECTIVE CONTINUA

1. *The Power Function*

The power function relating subjective magnitude (R) to stimulus magnitude (S) may be written

$$R = c(S + a)^n, \tag{1}$$

where c is concerned with the arbitrary unit of measurement, and a and n are constants to be experimentally determined. In this laboratory we have found reasons to prefer this form of the function to the usual form cS^n. The constant a may sometimes without serious error be omitted, but for small values of S it is not negligible and in some cases of shifting functions (2, 6, 14, 18) it may be important even for high values of S. The constant is easily solved

---

*Received in the Editorial Office on February 24, 1959, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This is one in a series of investigations made possible by a research grant from the Wallenberg Foundation. The present paper is partly based on unpublished research reports (4, 5, 7, 8, 9). Details which have been omitted here may be seen from the reports.

by the algebraic procedures developed by the present writer for treatment of fractionation data (6).

2. *The Computation of* $\triangle$ R

The experimentally obtained measure of sensitivity is usually in terms of S. It will be denoted by $\triangle S$, whatever method has been used to determine it. Our first problem is to transform this measure to the subjective continuum, i.e., to compute $\triangle R$.

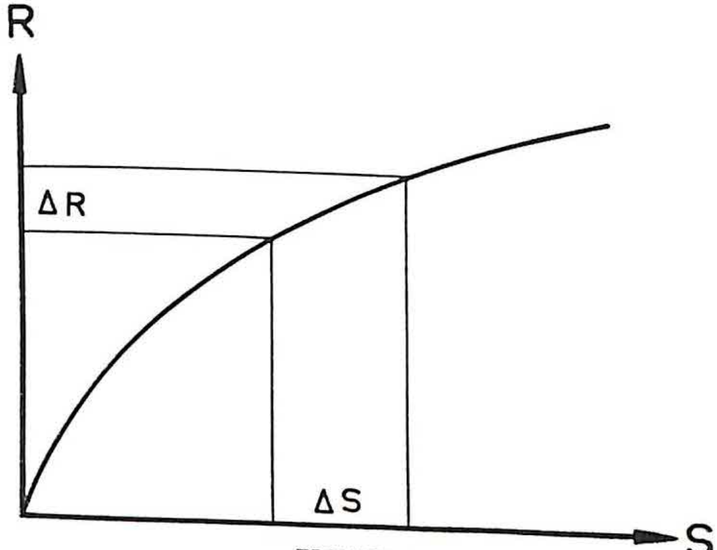


FIGURE 1
JUST NOTICEABLE DIFFERENCE ON THE STIMULUS CONTINUUM (S) AND ON THE SUBJECTIVE CONTINUUM (R)

Consider Figure 1. If $\triangle S$ is small compared to S, the corresponding segment of the curve is practically linear. Accordingly we have, approximately,

$$\triangle R = \frac{dR}{dS} \triangle S. \tag{2}$$

This relation was pointed out by the present writer in a previous paper, where dR/dS was determined by graphic procedures (3). When a mathematical function has been fitted to the plot of R against S, the derivative is found from this function. In the case of a power function

$$\frac{dR}{dS} = nc(S + a)^{n-1},$$

which may also be written

$$\frac{dR}{dS} = \frac{nR}{S+a}.$$

3. *Experimental Data*

Harper and Stevens (10), who established the veg scale, also obtained estimates of $\triangle R$, apparently by graphic means, on the basis of Oberlin's (14) measures of $\triangle S$ for *weight lifting*. They demonstrated that these estimates increased with S. In a previous paper by the present writer (3) Oberlin's data were converted into $\triangle R$ by Equation (2) and plotted against R as determined by Harper and Stevens; the general trend was linear with an R-intercept nearly 0. The present demonstration is based on the same data. They have been converted into $\triangle R$ by Equation (2) assuming a power function with the exponent 1.45 as averaged by Stevens and Galanter (17) from several investigations. The result is shown in Figure 2A. The trend of the plot is well approximated by

$$\triangle R = kR, \tag{3}$$

i.e., $\triangle R$ is directly proportional to R. This means that Weber's law is valid in the *subjective* continuum. It should be borne in mind that the data brought together in Figure 2A are not strictly comparable, in that the exponents of Oberlin's subjects may differ from the average value used here. Such a difference, however, would rather obscure than produce a simple relation.

In this laboratory, several scaling experiments have been carried out in recent years. In two of them measures of $\triangle S$ have also been obtained. These

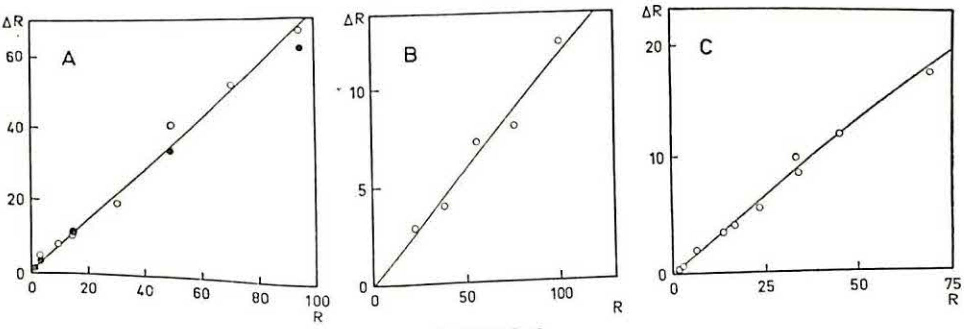


FIGURE 2

Just Noticeable Differences in Subjective Units, Plotted Against Subjective Magnitude

Diagram *A* shows values computed from Oberlin's data for lifted weights (Exp. I: empty circles; Exp. II: filled circles; one point from Exp. II is omitted because it is based on a somewhat different group of subjects). Diagram *B* shows results from an experiment on visually perceived velocity, and Diagram *C* results from an experiment on auditorily perceived time.

have been converted into $\triangle R$ by Equation (2) and are plotted against R in Figure 2B and C.

Diagram B shows results from an experiment on visually perceived subjective *velocity* (7), carried out with 10 subjects. The exponent of the power function was 1.77. The measure of $\triangle S$ was obtained by the method of reproduction (the variable stimulus was adjusted to appear equal to S in addition to the half-adjustments required by the fractionation method of scaling). Again Equation (3) is verified.

Diagram C shows results from an investigation with 10 subjects on auditorily perceived subjective *time* (8). This graph brings together estimates of $\triangle R$ according to Equation (2) from half- as well as equal- adjustments, since these two sets of data in this experiment yielded the same constant of proportionality k. (In the velocity experiment, k is higher for half-adjustments although the trend of proportionality remains the same.)

Finally, some recent data from an investigation on *pitch* will be shown. It was demonstrated that pitch was adequately described as a power function of frequency. The function exhibits a shift around 3 kc/s corresponding to $R = 35$ in the arbitrary unit of the experiment. These results have been reported elsewhere (2). The $\triangle R$ was computed from $\triangle S$ for half-adjustments and is plotted against R in Figure 3. There is a rather drastic shift about $R = 17$ (corresponding to $R/2 = 35/2$), so that we see two linear trends. In this case, however, no direct proportionality is seen. According to Stevens and his co-workers, pitch is a "metathetic" continuum, and $\triangle R$ would be expected to be constant over R (17). This does not seem to be the case in our experiment, which may be stated without any further comments in the present context.

4. *Conclusion*

It has been shown that $\triangle R$ is approximately proportional to R in three subjective continua: weight, visually perceived velocity, and auditorily perceived time. The result obtained means that Weber's law holds true in these *subjective* continua, and this may also be true for other "prothetic" continua. Our next problem will be to investigate the adequate form of this law when applied to the stimulus continuum for which it was originally formulated.

C. DIFFERENTIAL SENSITIVITY IN TERMS OF STIMULATION

1. *Derivation of Equations*

Solving the simultaneous Equations (1), (2), and (3) for $\triangle S$, we obtain

$$\triangle S = \frac{k}{n}(S + a), \tag{4}$$

and for the Weber fraction we have the equation

$$\frac{\triangle S}{S} = A + B/S, \tag{5}$$

where $A = k/n$, and $B = ka/n$.

It is seen that the Weber fraction will decrease with increasing S, if $a > 0$. This corresponds to well known experimental facts; only for moderate and high values of S does the fraction converge towards an approximately constant value.

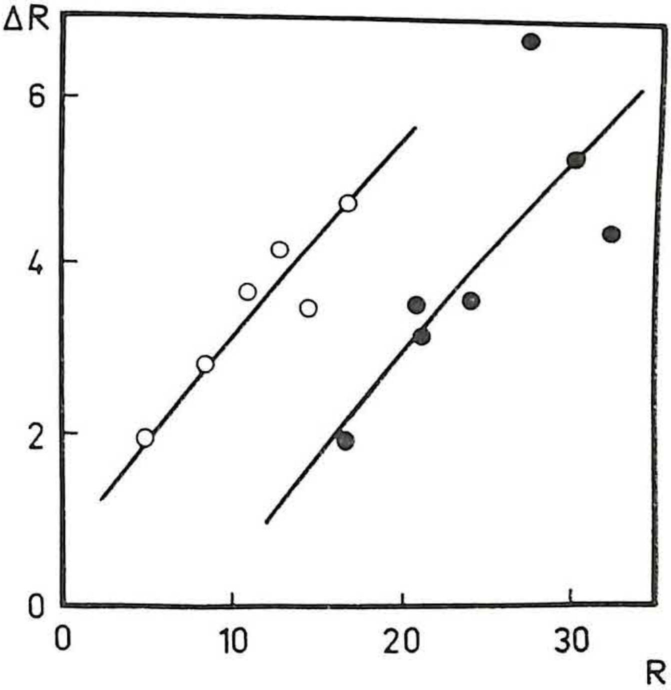


FIGURE 3
JUST NOTICEABLE DIFFERENCES ON THE PITCH CONTINUUM, COMPUTED FROM HALF-ADJUSTMENTS

Two linear trends are seen, but no simple proportionality. The shift corresponds to a shift in the R-S function; see text.

The Weber fraction, however, is constant if $(S + a)$ is chosen as denominator instead of S,

$$\frac{\triangle S}{S + a} = A, \tag{6}$$

where $A = k/n$ as before. This form of the Weber fraction has long been recognized as a better approximation to experimental data than $\triangle S/S$. The purpose of the derivation in this paper is to show that this formula may be considered a *semi-rational* equation, derived as it is from the two empirical Equations (1) and (3) linked together by means of the rational operation indicated by Equation (2).

2. *Applications to Experimental Data*

Data concerning visual intensity discrimination have been carefully examined by Hecht (11). Two of these sets of data, those of Blanchard (1) and Lowry (12), are shown in Figure 4. As pointed out by Hecht, the data fall into two sections indicative of rod and cone functions. The constants A and a of Equation (6) were determined from linear plots of $\triangle S$ against

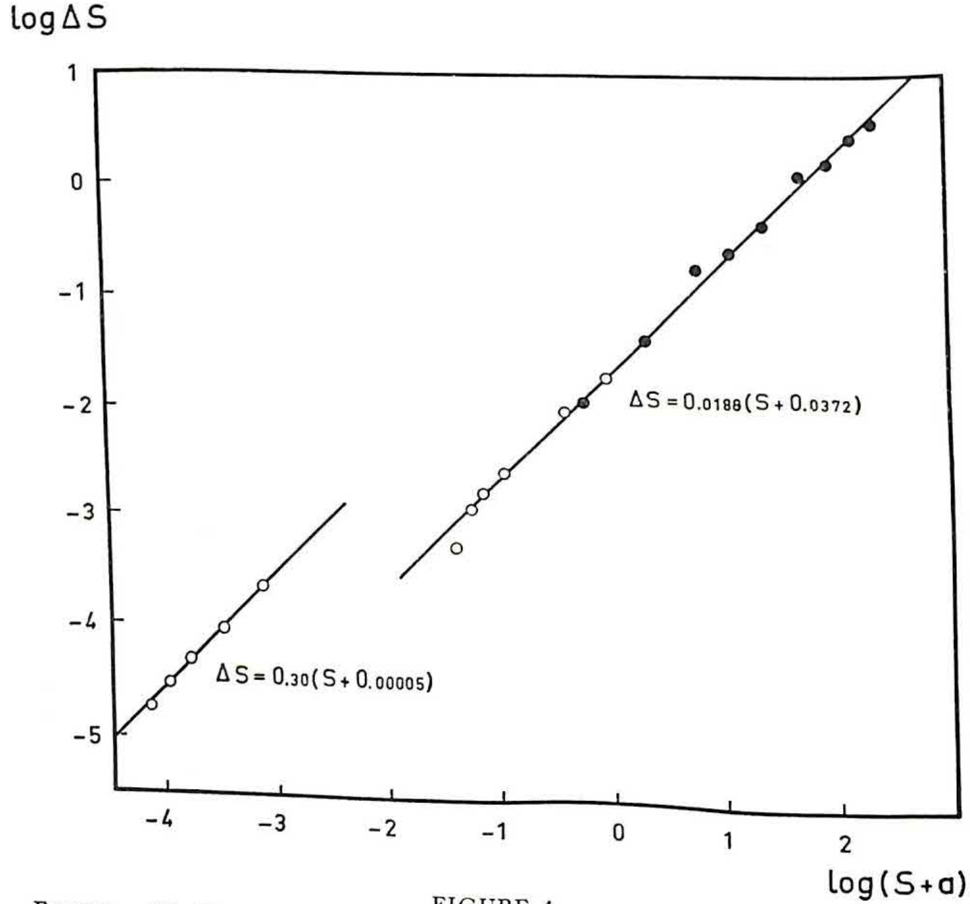


FIGURE 4
Equation (6) Fitted to Data (Log Millilamberts) from Scotopic and Photopic Vision Obtained by Blanchard (Empty Circles) and Lowry (Filled Circles)

S. Equation (6) is fitted to the data in the logarithmic plot of Figure 4. The fit is quite good for both scotopic and photopic vision. This is indirect evidence that $\triangle R$ is directly proportional to R also in the case of brightness, although the evidence is less convincing than that obtained from a direct approach to the problem, such as reported in the preceding section.

Experimental data on differential intensity sensitivity for *pure tones* have been published by Riesz (15). Equation (6) does *not* represent these data adequately. As has been previously shown by the present writer, however, it is possible to derive a more general form of Equation (6) by assuming $\triangle R$ to be a power function of R, and this equation yields a perfect fit to the data of Riesz (4). A direct investigation of the differential sensitivity in the loudness continuum, however, will yield more reliable information.

D. Absolute and Differential Sensitivity

The constant a plays a decisive rôle in the formulae derived above for the Weber fraction. The history of this constant within our laboratory will be briefly reviewed.

A few years ago we wrote the power function in the form

$$R = c(S - S_0)^n,$$

and first supposed S_0 to be related to the absolute threshold. By means of our method for treating fractionation data, we were able to obtain estimates of S_0. In one experiment after another it turned out to be *negative*. This means that $R > 0$, when $S = 0$. This is the reason why in this paper S_0 is replaced by the symbol a which is supposed to be positive in most experiments.

One tentative interpretation of a is as follows. Even if (external) stimulation is 0, there is probably in general a certain amount of *sensory noise,* resulting from spontaneous nervous activity. This would explain why $R > 0$, even when $S = 0$. This fact may be represented by Equation (1) when a has a small positive value. It is illustrated by the power function in Figure 5.

This interpretation might suggest another form of the power function,

$$R = R_0 + S^n,$$

where R_0 is the basic perceptual noise, to which the subjective magnitude produced by external stimulation is *added*. This form of the function is mathematically not identical with that of Equation (1), but at present it seems hardly possible to make a choice between the two functions on the basis of experimental work; both of them will fit about equally well to experimental data.

Our interpretation of a, however, again draws attention to the concept of absolute thresholds. If $R > 0$ without any external stimulation, then the

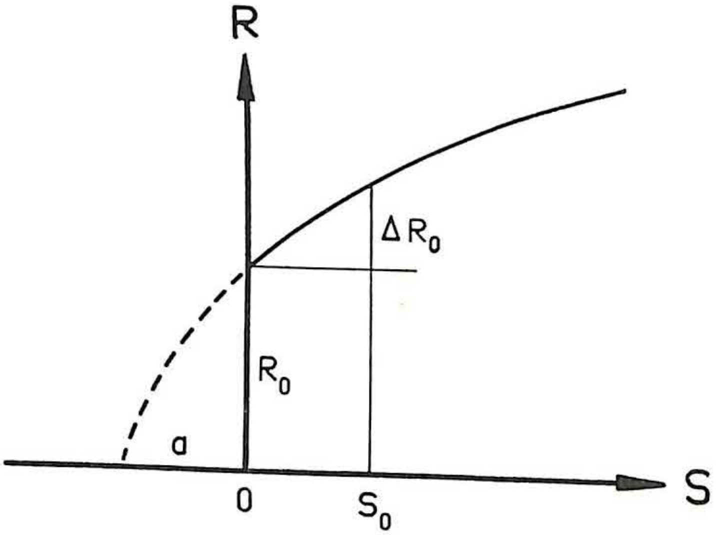


FIGURE 5
Possible Relations Between Absolute and Differential Sensitivity

subjective correlate of *liminal* stimulation (S_0 in the usual sense) is a just noticeable *increase* on R. Assuming Equation (4) to be valid also in this liminal region,

$$S_0 = \frac{k}{n} a.$$

Substituting in Equation (4) we obtain

$$\frac{\triangle S}{S_0} = \frac{S + a}{a},$$

which may be written in the simple form

$$\triangle S = S_0 S', \tag{7}$$

where S' is measured from $-a$ with a as the unit.

This equation serves mainly the purpose of illustrating, in principle, how absolute and differential sensitivity as well as the relation between stimulus magnitude and subjective magnitude might be treated within a common theoretical framework.

E. Summary

1. Discriminal sensitivity is measured in subjective terms ($\triangle R$) by proper transformations of just noticeable differences and similar measures

($\triangle S$). It is found that $\triangle R$ is directly proportional to R in three subjective continua investigated: weight, visually perceived velocity, and auditory time. This means that the Weber law in its original form applies to these subjective continua. In addition, two distinct linear trends are found in the "metathetic" continuum of pitch.

2. On the basis of the experimentally verified proportionality, the adequate form of Weber's law in terms of stimulation is derived. It is fitted to experimental data, and the fit is indirect evidence of the same proportionality with regard to the subjective continua of scotopic and photopic brightness.

3. Finally it is shown in principle how the concept of absolute sensitivity may be considered a special case of differential sensitivity.

REFERENCES

1. BLANCHARD, J. The brightness sensibility of the retina. *Phys. Rev.*, 1918, **11**, 81-99.
2. EISLER, H., & EKMAN, G. A mechanism of subjective similarity. *Acta Psychol.*, 1959, **15** (In press).
3. EKMAN, G. Discriminal sensitivity on the subjective continuum. *Acta Psychol.*, 1956, **12**, 15-18.
4. ———. A note on generalized psychophysical laws. Psychol. Lab. Univ. Stockholm, April, 1956, No. 32. (Unpublished.)
5. ———. Three methods of estimating discriminal dispersion. Psychol. Lab. Univ. Stockholm, October, 1957, No. 50. (Unpublished.)
6. ———. Two generalized ratio scaling methods. *J. of Psychol.*, 1958, **45**, 287-295.
7. EKMAN, G., & DAHLBÄCK, B. A subjective scale of velocity. Psychol. Lab. Univ. Stockholm, February, 1956, No. 31. (Unpublished.)
8. EKMAN, G., & FRANKENHAEUSER, M. Subjective time scales. Psychol. Lab. Univ. Stockholm, September, 1957, No. 49. (Unpublished.)
9. EKMAN, G., & KÜNNAPAS, T. Subjective dispersion and the Weber fraction. Psychol. Lab. Univ. Stockholm, January, 1957, No. 41. (Unpublished.)
10. HARPER, R. S., & STEVENS, S. S. A psychological scale of weight and a formula for its derivation. *Amer. J. Psychol.*, 1948, **61**, 343-351.
11. HECHT, S. A theory of visual intensity discrimination. *J. Gen. Physiol.*, 1935, **18**, 767-789.
12. LOWRY, E. M. The photometric sensibility of the eye and the precision of photometric observations. *Opt. Soc. Amer.*, 1931, **21**, 132-136.
13. OBERLIN, K. W. Variations in intensitive sensitivity to lifted weights. *J. Exp. Psychol.*, 1936, **19**, 438-455.
14. REESE, T. W. The application of the theory of physical measurement to the measurement of psychological magnitudes with three experimental examples. *Psychol. Monog.*, 1943, **55**, No. 251.
15. RIESZ, R. R. Differential intensity sensitivity of the ear for pure tones. *Phys. Rev.*, 1928, **31**, 867-875.
16. STEVENS, S. S. On the psychophysical law. *Psychol. Rev.*, 1957, **64**, 153-181.
17. STEVENS, S. S., & GALANTER, E. H. Ratio scales and category scales for a dozen perceptual continua. *J. Exp. Psychol.*, 1957, **54**, 377-411.

18. Taves, E. H. Two mechanisms for the perception of visual numerousness. *Arch. of Psychol.*, 1941, **37**, No. 265.

Psychological Laboratory
University of Stockholm
Teknologgatan 8
Stockholm Va, Sweden